



**A University of Sussex DPhil thesis**

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

UNIVERSITY OF SUSSEX

# Applying Pause Analysis to Explore Cognitive Processes in the Copying of Sentences by Second Language Users

---

**PUTRI AFZAN MARIA ZULKIFLI**

**2013**

---

School of Engineering & Informatics

---

DISSERTATION SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS  
OF THE DEGREE OF DOCTOR OF PHILOSOPHY

To my husband, Faizul Iqmal.  
To my parents, Zulkifli and Tengku Muhaizah.  
To my parents-in-law, Mohd Kamil and Zulfah.  
To my brothers and sisters, Nadira, Faizatul, Faiz, Syafiq and Aman.  
To my children, Farish, Imran and Matin.

Thank you for supporting and believing in me.

## ACKNOWLEDGEMENT

First of all I would like to thank ALLAH Almighty, Who gave me the courage, health, and energy to accomplish my thesis in due time, and without Whose help it would not have been possible to complete this study, which required untiring efforts, within the time limits.

Motivation, encouragement, guidance, corrections, advice, and overall support are the key elements required from a supervisor to allow someone to complete a thesis to a good standard and quality within the deadline. It is a matter of utmost pleasure for me to extend my gratitude and give due credit to my supervisor, Prof. Peter Cheng, whose support has always been there in time of need and who provided me with all these key elements so I was able to complete my dissertation within the time frame.

Acknowledgement would be incomplete without extending my gratitude to everyone who has given full support to my completion of this thesis. Very special thanks to my husband, Faizul Iqmal, who has always been my backbone, and to my children, Farish, Imran and Matin, for giving me hope and inspiration. I am also very grateful to both my parents, Tengku Muhaizah Putra (my mother) and Zulkifli Abu Bakar (my father) for sending me abroad in the first place to learn how to be independent and to appreciate that living is not as easy as we thought it was. I have learned so much through living overseas. I thank my parents-in-law, Mohd Kamil (my father-in-law) and Zulfah Zainuddin (my mother-in-law) for also being very supportive throughout my studies. I thank all my friends in Leeds and Manchester, especially Evi Indrasari and family, Syarifah Fadliyatun (Kak Aton) and Jaja Shahril and family, for being so supportive. I thank all my lab mates at Sussex University: Unaizah (who has always been there through my ups and downs), Grecia (thank you so much for all your positive words), Izzah (thank you for being very positive too), Rossano, and Ronald, for all their support. I thank all my neighbours who were so friendly and helpful whilst we were living in Park Village: Tayyeb, Mahmoud and family, Zainab and family, Gazela and family, Waleed Al-Hasibi and family, Abid Masoud and family, Bilal and family, and of course everyone else whom I could not list here.

My thanks are also due to my examiners, Kristyan Spelman Miller and Bill Keller, whose valuable comments and suggestions made a colossal contribution in improving my dissertation.

Last but not least, I extend my thanks to everyone whose names I have not mentioned. I thank everyone for their moral support and prayers for my health and successful completion of my dissertation.

Final gratitude goes to the Ministry of Higher Education Malaysia (MOHE) and Universiti Pendidikan Sultan Idris (UPSI) for their crucial financial support, without which this thesis would never have been completed.

**2013**

**Putri Afzan Maria Zulkifli**

## **Applying Pause Analysis to Explore Cognitive Processes in the Copying of Sentences by Second Language Users**

Putri Afzan Maria Zulkifli  
School of Informatics  
University of Sussex  
Brighton  
afzanmaria@gmail.com

### **ABSTRACT**

Pause analysis is a method that investigates processes of writing by measuring the amount of time between pen strokes. It provides the field of second language studies with a means to explore the cognitive processes underpinning the nature of writing. This study examined the potential of using free handwritten copying of sentences as a means of investigating components of the cognitive processes of adults who have English as their Second Language (ESL).

A series of one pilot and three experiments investigated possible measures of language skill and the factors that influence the quality of the measures. The pilot study, with five participants of varying English competence, identified copying without pre-reading to be an effective task and 'median' at the beginning of words to be an effective measure. Experiment 1 (n=20 Malaysian speakers) found jumbled sentences at the letter and word levels to effectively differentiate test-taker competence in relation to grammatical knowledge. Experiment 2 (n=20 Spanish speakers) investigated the jumbling effects further, but found that participants varied their strategy depending on the order of the sentence types. As a result, Experiment 3 (n= 24 Malaysian speakers) used specific task instructions to control participant strategy use, so that they either attended to the meaning of the sentences, or merely copied as quickly as possible. Overall, these experiments show that it is feasible to apply pause analysis to cognitively investigate both grammar and vocabulary components of language processing.

Further, a theoretical information processing model of copying (MoC) was developed. The model assists in the analysis and description of (1) the flow of copying processes; (2) the factors that might affect longer or shorter pauses amongst participants of varying competence level; and (3) sentence stimuli design.

## Contents

DECLARATION .....	2
ACKNOWLEDGEMENT .....	4
ABSTRACT .....	6
CHAPTER 1 Introduction: What is Pause Analysis & How Can It Be Used? .....	15
1.1 How Can Copying be used to study the Underlying Cognitive Processes of Language Components Captured? .....	16
1.2 What Kinds of Stimuli are Used? .....	18
1.3 What Constitutes a Pause? .....	19
1.4 What are the Instruments Used to Measure Pauses? .....	22
1.5 How Can Pauses be Used to Investigate the Components of Cognitive Processes in Language Processing of ESLs? .....	24
1.6 How do you Validate the Measure? .....	25
1.7 Why focus on Second Language Learners? .....	27
1.8 An Example of How the Cognitive Processes in Language Processing in ESLs is Measured.. .....	28
1.9 What are the Underlying Cognitive Processes Involved? .....	30
1.10 Conclusion .....	31
CHAPTER 2 Literature Review: Studies of Pauses, Sentence-Copying & Second Language Processing .....	33
2.1 Introduction .....	33
2.2 Pause Analysis: An Alternative Method .....	33
2.2.1 The Initial Approach .....	34
2.2.2 The Role of Computers & Technology .....	35
2.2.3 Pauses as Reflections of Cognitive Processes .....	39
2.3 Copying: An Insight into Understanding the Cognitive Processes .....	40
2.3.1 Why Copying? .....	41
2.3.2 Previous Research in Copying .....	42
2.3.3 The Process of Copying & Working Memory .....	43
2.3.4 Sentence-Copying as a Predictive Task to Investigate the Cognitive Processes in Language Processing of ESLs .....	44
2.4 Second Language Processing .....	46
2.4.1 A Glimpse of the Second Language Learner - How Learner Acquire SL Knowledge Impacts on the Cognitive Processes? .....	47
2.4.2 A Glimpse of Beginning SL Reading .....	47

2.4.3	Why: Learner Differences? .....	49
2.4.4	How?: The Learning Processes.....	51
2.5	Factors Affecting the Cognitive Demands of Copying: The Processing Costs.....	53
2.5.1	Working Memory Related Processes .....	54
2.5.2	Handwriting Related Processes.....	57
2.5.3	Word Related Processes .....	61
2.5.4	Meaning Related Processes .....	66
2.5.5	The Effects of Being an English as a Second Language Learner (ESL) .....	69
2.5.6	Summary of Factors that Affect the Cognitive Demand of Copying.....	72
2.6	Overall Conclusion .....	72
CHAPTER 3	Pilot Study: The Suitability of Pause Analysis as a Tool for Measuring the Cognitive Processes of Language Processing in Sentence-Copying.....	74
3.1	Introduction .....	74
3.2	The Sentence Stimuli Design.....	77
3.2.1	The Baseline .....	77
3.2.2	The Sentences .....	79
3.3	Method .....	84
3.3.1	Participants .....	84
3.3.2	Apparatus.....	85
3.3.3	Procedure.....	85
3.3.4	Overall Structure of this Experiment: The Variables.....	86
3.3.5	Analysis .....	87
3.4	Results.....	87
3.4.1	Suitable Method of Copying Sentences & Measures of Cognitive Processes in Language Processing.....	87
3.4.2	The Distribution of Pauses in Different Types of Sentences.....	88
3.4.3	The Baselines.....	92
3.5	Discussion .....	94
3.5.1	What is the Suitable Method for Copying: IC or IR? .....	95
3.5.2	What is the Effective Measure for Cognitive Processes in Language Processing? ...	97
3.5.3	What is the Appropriate Sentence Type for the Purposes of this Thesis? .....	99
3.5.4	What is the Suitable Baseline Stimulus? .....	99
3.6	Conclusion.....	100
CHAPTER 4	A Theoretical Model of Copying (MoC).....	102
4.1	Introduction .....	102
4.2	Background .....	104



4.2.1	The Levels of Processing .....	104
4.2.2	Word Characteristics.....	106
4.3	Model Structure.....	110
4.3.1	The Components of the Model .....	111
4.3.2	The Flowchart .....	113
4.3.3	The Routes & Processing Load Prediction .....	114
4.3.4	Route Predictions.....	117
4.4	General Simulation of the Model .....	119
4.4.1	Word Group Level Chunking .....	120
4.4.2	Word Level Chunking .....	120
4.4.3	Syllable Level Chunking.....	121
4.5	Conclusion.....	128
CHAPTER 5	Experiment 1: Investigating Suitable Sentence Stimuli to Assess the Cognitive Processes of Language Processing.....	130
5.1	Introduction .....	130
5.2	Method .....	132
5.2.1	Participants .....	132
5.2.2	Apparatus.....	132
5.2.3	Procedure.....	133
5.2.4	The Stimuli: Sentences & Baselines .....	133
5.2.5	The Independent Measure of English Language.....	133
5.2.6	The Overall Structure of this Experiment .....	134
5.2.7	Analysis .....	135
5.3	Results.....	135
5.3.1	Online Oxford Diagnostic English Test Level.....	135
5.3.2	Language Performance versus Overall Pause Value .....	136
5.3.3	Language Performance Versus Pause Levels .....	137
5.3.4	Normalisation: Correcting Individual Writing Speed .....	138
5.3.5	Refining the Tasks .....	141
5.4	Discussion .....	143
5.4.1	Confirming the Pilot Findings.....	144
5.4.2	Which Sentence Stimulus is Effective for Measuring Language? .....	144
5.4.3	The Role of Pauses in Obtaining Information on Language Processing.....	146
5.4.4	The Implications of the MoC.....	149
5.5	Conclusion.....	150

CHAPTER 6	Experiment 2: Investigating a Suitable Jumbling Condition to Assess Components of Language Processing.....	152
6.1	Introduction .....	152
6.2	The Sentence Stimuli Design.....	153
6.2.1	Why Jumbling Conditions?.....	154
6.2.2	Levels of Jumbling .....	154
6.2.3	The Design of Stimuli .....	158
6.2.4	Making Simulations with the Jumbling Conditions.....	159
6.3	Method .....	170
6.3.1	Participants .....	170
6.3.2	Apparatus.....	170
6.3.3	Procedure.....	170
6.3.4	Language Performance Test .....	172
6.3.5	The Overall Structure of this Experiment: The Variables.....	173
6.3.6	Analysis .....	174
6.4	Results.....	174
6.4.1	Validating the Independent Language Tests: ODT & VST .....	175
6.4.2	The General Effects of Copying Different Jumbling Conditions.....	175
6.4.3	Data Aggregated According to the Jumbling Conditions & First Jumbling Type Received .....	176
6.4.4	Data Aggregated by Sentence Set, Focusing Only on Sentence Set 1 .....	178
6.4.5	Normalisation.....	178
6.5	Discussion .....	180
6.5.1	Are ODT and VST Measuring Different Components? .....	182
6.5.2	Is Jumbling a Good Way to Design Stimuli? .....	182
6.5.3	Did the Normalisation Appear to Improve Measures? .....	183
6.5.4	Why Did Both Words Jumbled and Letters Jumbled Not Work? .....	184
6.5.5	What is Happening During Copying in the First & Second Trials? .....	185
6.5.6	Limitation of the Experiment .....	189
6.5.7	Implications of the MoC.....	189
6.6	Conclusion.....	191
CHAPTER 7	Experiment 3: Investigating the Effects of Task Type (Processing Conditions) on the Strategy of Copying.....	192
7.1	Introduction .....	192
7.2	The Sentence Stimuli Design.....	194
7.2.1	The Baselines.....	194

7.2.2	The Sentences .....	194
7.2.3	The Processing (Task Type) Conditions .....	195
7.3	Method .....	196
7.3.1	Participants .....	196
7.3.2	Apparatus .....	196
7.3.3	Procedure .....	196
7.3.4	The Language Tests .....	198
7.3.5	Overall Structure of this Experiment: The Variables .....	198
7.3.6	Analysis .....	199
7.4	Results .....	200
7.4.1	General Language Competency Scores: ODT, VST & Pauses .....	200
7.4.2	The True/False Questions .....	201
7.4.3	The Effects of Task Type (Meaning & Fast-copying) at General Level .....	202
7.4.4	The Effects of the Meaning Task Type during the First Trial .....	206
7.4.5	Confirming the Effects of the Fast-copying Condition in the First Trial .....	208
7.4.6	Findings on Normalisation .....	209
7.5	Discussion .....	210
7.5.1	Did the Task Type – Meaning or Fast-copying – Control Copying Strategy? .....	210
7.5.2	What is the Effective Measure? .....	211
7.5.3	What is a Good Stimulus? .....	212
7.5.4	What is the Suitability of the True or False Questions? .....	213
7.5.5	What is the Effect of Normalisation? .....	214
7.5.6	Implications of the MoC .....	214
7.6	Conclusion .....	215
CHAPTER 8	The Experiments Revisited .....	218
8.1	Introduction .....	218
8.2	Experiment 1 Revisited .....	218
8.2.1	Results .....	219
8.2.2	Discussion for Experiment 1 Revisited .....	228
8.3	Experiment 2 Revisited: Comparing Methods A & B in the First Trial .....	230
8.3.1	Results .....	230
8.3.2	Discussion for Experiment 2 Revisited .....	231
8.4	Overall Discussion & Conclusion .....	232
CHAPTER 9	Discussion & Conclusion .....	234
9.1	Introduction .....	234

9.2	Discussions of Findings, Limitations & Implications on the Future Work.....	235
9.2.1	What is the Best Type of Copying? .....	235
9.2.2	What Kind of Sentence Stimuli?.....	237
9.2.3	Can Components of Language be Distinguished?.....	238
9.2.4	What Kind of Measure is Suitable? .....	238
9.2.5	Does Data Need to be Normalised?.....	240
9.2.6	The Model of Copying .....	240
9.2.7	Critical Issues.....	241
9.3	Overall Conclusion .....	243
Appendices	.....	245
	Baseline Writing for Chapter 4 and 5.....	245
	Baseline Writing for Chapter 6 and 7.....	245
	Sentence Stimuli for Chapter 4 and 5 .....	246
	Sentence Stimuli for Chapter 6.....	247
	Sentence Stimuli for Chapter 7 .....	250
References	.....	253

## Figures

Figure 1.1 Illustration of the pause levels used in this study .....	20
Figure 1.2 The writing stimulus .....	23
Figure 1.3 An example of the written stimuli captured and extracted in TRACE .....	23
Figure 1.4 An example of how pauses differ between the most and least competent within a group of participants .....	29
Figure 2.1 Handwriting Production Model by van Galen (1991) .....	59
Figure 2.2 Basic Architecture of the Extended Dual Route Cascaded Model (dotted lines show the extensions as compared with the standard DRC model) (Verhoeven et al., 2006) .....	65
Figure 3.1 The Baseline Stimuli .....	77
Figure 3.2 The overall structure of this experiment.....	86
Figure 3.3 A close-up view of L0 pause levels .....	90
Figure 3.4 A close-up view of the L1 pause levels, demonstrating the predicted preserved rank order .....	90
Figure 3.5 A close-up view of L2 pause levels .....	91
Figure 3.6 An overview of Pause Distribution at L0, L1 and L2 pause levels across all participants and sentence types.....	92
Figure 3.7 Baseline L1 (Median) Pauses correlated against the Language Competence Rank Order .....	93
Figure 3.8 Correlation values based on median of L1 and L2 pauses between the baselines of lower case name writing against the sentence stimuli .....	94
Figure 3.9 An example of L1 pause distribution for the most competent participant (A) in this experiment .....	98
Figure 3.10 An example of L1 pause distribution for the least competent participant (T) in this experiment .....	98
Figure 4.1 Illustration of the pause levels (in brackets) and the levels of processing .....	104
Figure 4.2 Example of chunking in relation to the levels of processing and pause levels .....	105
Figure 4.3 Theoretical MoC (Model of Copying) .....	110
Figure 4.4 The flowchart explaining the decision making processes involved in sentence-copying .....	114
Figure 4.5 Example of chunking at group of words level and the processes that might take place during the sentence-copying.....	120
Figure 4.6 Example of chunking at individual word level and the processes that might take place during the sentence-copying.....	120
Figure 4.7 Example of the assumed chunking at syllable level and the processes that might take place during the sentence-copying.....	121
Figure 4.8 Example of the assumed chunking at syllable level and the processes that might take place during the sentence-copying.....	121
Figure 4.9 Details of the features in Figures 3.10.....	123
Figure 4.10 Examples of PLP Plotting.....	125
Figure 4.11 Examples of real data and PLP data plotted on a line graph.....	126
Figure 4.12 An illustration of the pause distribution differences between MC and LC that the thesis aims to uncover.....	127
Figure 5.1 The overall structure of this experiment.....	134
Figure 5.2 Baselines correlated against the sentence stimuli at Letter (L1) Level .....	139
Figure 5.3 Baselines correlated against the sentence stimuli at Word (L2) Level .....	140
Figure 5.4 An example of pause distribution detail for sentence number 24, to observe the occurrences of long pauses that can be seen between the most competent (solid red line) and least competent (dotted blue line) participants .....	148
Figure 6.1 2x2 Factorial Design of the Experiment.....	155
Figure 6.2 Sentence example for no jumbling condition.....	163
Figure 6.3 PLP values for a no jumbling condition sentence .....	163
Figure 6.4 Sentence example for words jumbled condition .....	164
Figure 6.5 PLP values for a words jumbled condition sentence .....	165
Figure 6.6 Sentence example for a letters jumbled condition.....	166

Figure 6.7 PLP values for a letters jumbled condition sentence .....	167
Figure 6.8 Sentence example for both jumbled condition .....	168
Figure 6.9 PLP values for a both jumbled condition sentence .....	168
Figure 6.10 The order of stimuli .....	170
Figure 6.11 The design of the stimuli order .....	171
Figure 6.12 The overall structure of this experiment .....	173
Figure 6.13 The correlation between ODT and VST .....	175
Figure 6.14 Overall results (all sentence sets) of copying the different jumbling conditions, aggregated by the first jumbling type received .....	177
Figure 6.15 The results from copying the first sentence set, aggregated by the first jumbling type received .....	179
Figure 6.16 The normalised results of the first sentence set, aggregated by the first jumbling type received .....	180
Figure 7.1 Sentence order as given to the participants .....	197
Figure 7.2 Equal distributions of sentence stimuli and jumbling type .....	197
Figure 7.3 Example of True or False questions given to participants .....	198
Figure 7.4 The overall structure of this experiment .....	199
Figure 7.5 Relationships between ODT and VST .....	200
Figure 7.6 Results showing the effects of Task Type Order and Sentence Presentation on ODT with L2L1 pause data .....	203
Figure 7.7 ODT Results at all pause levels according to the order of sentence presentation .....	204
Figure 7.8 Results showing the Effects of Task Type Order and Sentence Presentation on VST with L2L1 pause data .....	204
Figure 7.9 Correlations of the first trial (L2L1) for each jumbling type (6 participants each) against ODT-VST for each measure .....	207
Figure 7.10 Normalisation conducted only on the First Trial .....	210
Figure 8.1 An example of pause distribution for Set I .....	222
Figure 8.2 An example of pause distribution for Set II .....	223
Figure 8.3 An example of pause distribution for Set III .....	224
Figure 8.4 Example of Group I (sentence 19), where chunking can be mostly seen at L2 (word) level .....	226
Figure 8.5 Example of Group II (sentence 21), where chunking can be seen at both L2 (word) and L1 (syllable) level .....	227
Figure 8.6 Example of Group III (sentence 24), where chunking can be seen mostly at L1 (syllable) level .....	228
Figure 8.7 The first trial results of Methods A and B for Experiment 2 data .....	231

## Tables

Table 2.1 Writing systems of the world .....	48
Table 2.2 The order of syntax structure according to language .....	49
Table 3.1 The Sentence Stimuli .....	79
Table 3.2 Participants' Backgrounds.....	85
Table 3.3 Correlation values calculated for each pause level and type of sentence (*p<.05; **p<.01; ***p<.001).....	89
Table 4.1 The labelling used to describe the word characteristics .....	108
Table 4.2 The functions of each process in the model .....	112
Table 4.3 The Processing Load Prediction (PLP) .....	116
Table 4.4 The possibilities of each word characteristic combination to be processed at specific chunking level .....	117
Table 4.5 Median of PLPs for the three examples .....	124
Table 5.1 The correlation values between the three tests of ODT (*p<.001) .....	136
Table 5.2 Correlations between the Overall Pauses of L2, L1 and L0 against the ODT tests (*p<.05, **p<.01 and ***p<.001) .....	136
Table 5.3 Correlations between ODT scores and pause durations across writing levels (*p<.01 and **p<.001).....	137
Table 5.4 Summary of all correlations between L2 and L1 pauses against ODT and Sentence Stimuli (*p<.05; **p<.01; ***p<.001).....	138
Table 5.5 Correlations between ODT scores and the Pause Quotient and Pause Difference (*p<.05; **p<.01; ***p<.001) .....	141
Table 5.6 Correlations for each sentence stimuli with Medians of L2s and Normalised data in search of the effective sentence (*p<.05, **p<.01 and ***p<.001) .....	142
Table 5.7 Correlations between ODT Advanced and the refined pauses, grouped by the correlation value, before and after normalisation (*p<.01 and **p<.001) .....	143
Table 6.1 Word Characteristics.....	160
Table 6.2 Predictions of possible processes occurring during copying for the extreme conditions of Most Competent and Least Competent participant .....	161
Table 6.3 Median of PLP results from processing the sample sentences .....	169
Table 6.4 Trial order number for each participant.....	172
Table 6.5 Correlations between the English tests and pauses (*p<.05).....	176
Table 7.1 Correlations performed with three measures and three combinations of pauses (*p<.05) .....	201
Table 7.2 Number of correct responses based on the jumbling type received .....	201
Table 7.3 The number of incorrect responses based on the jumbling type and the order in which they were received .....	201
Table 7.4 Correlations of the first trial for ODT and VST with different measures, grouped according to jumbling type (*p<.05).....	207
Table 7.5 Correlation values between the fast-copying condition at first trial and NW (N=6, N.S) .....	208
Table 7.6 Correlation values between the meaning condition at first trial and NW.....	209
Table 8.1 Comparison of correlations between pauses and language performance using Methods A and B (*p<.05; **p<.01; and ***p<.001).....	219
Table 8.2 Correlations of Method A, Method B and Means of Pauses at different levels (*p<.05; **p<.01; ***p<.001).....	220

## CHAPTER 1 Introduction: What is Pause Analysis & How Can It Be Used?

***Pause analysis*** is a method that uses temporal aspects to represent the processes taking place during the activity of writing. Back in the 1980s, Matsuhashi (1981, 1982 & 1987) was one of the first researchers to explore real-time testing methods in relation to writing composition, such as think-aloud protocols and video capture of writing activities. This research by Matsuhashi opened a new perspective into cognitive processing. The method enables an understanding of what is occurring during the cognitive process of writing. This has led to the development of the *online* method of ***pause analysis***.

First of all, *pause analysis* has already been used in studies of writing such as the studies of syllable processing in words (e.g. Kandel et al., 2006a & 2006b) and writing problems in dyslexia (Van Genuchten & Cheng, 2009). This innovative approach to language assessment is made possible by acquiring rich raw data of the processes that occur during the activity of writing. Through the richness of these data, a greater amount of processing is revealed from participants of different levels of competence in the language, reflected in the different pause lengths at certain points within the writing processes. These pauses act as signals of cognitive processes. Schilperoord (1996) reported that *'longer pauses reflect cognitive processes that are relatively more effortful compared to processes reflected by shorter pauses'*. This means the pauses might reflect some important processes occurring: a longer pause may occur due to the effort to comprehend the meaning of a word, whereas a shorter pause is produced in processing a word that is familiar, because of the automaticity of producing the word without semantically processing it. Very rich data can be captured using this method, and the way in which it is able to portray the cognitive aspects of language processing may represent a novel approach in being used for measuring components of language competence.

Researchers have applied *pause analysis* in a number of studies involving different tasks, including: text production (e.g. Schilperoord, 1996, 2001; Torrance & Jeffery, 1998; Spelman Miller, 2000a, 2000b, 2006; Spelman Miller & Sullivan, 2006; Wengelin, 2006); the drawing of simple geometric patterns (Cheng, McFadzean & Copeland, 2001); the writing of number sequences (Cheng & Rojas-Anaya, 2005); the writing of familiar and unfamiliar words (Cheng & Rojas-Anaya, 2006); the copying of mathematical equations by experts and novices (Cheng & Rojas-Anaya, 2007); schema



and chunk production in drawings (Obaidallah & Cheng, 2009); writing from memory by dyslexic children (van Genuchten, Cheng, Leseman & Messer, 2009); writing multiple sentences (van Genuchten & Cheng, 2010); and finally, the studies of children's copying (e.g. Grabowski et al., 2010). All these studies have established the usefulness of the method. Given the potential of using pause analysis as an online method, this study intends to explore whether such an approach is feasible as a means to assess components cognitive processes in language processing in second language learners.

Pause analysis as an online approach offers many advantages. It can reduce the time needed for the study: both the time spent conducting the analysis and calculating scorings, and the time spent for participants to carry out the test. It may also be more cost effective, because there is no need for a manual coder (i.e. the human manual judgment of assessment items). It may also offer a more precise and accurate measure, which increases the quality of the measure. The online method could also be more precise in assessing participants, because of the rich raw data from which measures can be derived.

### **1.1 How Can Copying be used to study the Underlying Cognitive Processes of Language Components Captured?**

In order to investigate whether pause analysis can be used to measure the underlying cognitive processes of language components, this research firstly needs to investigate what kind of writing tasks is able to expose the underlying cognitive processes of the language components. Writing tasks are diverse and may include text production in essay compositions and the writing of poems or short stories. However, these tasks focus on the higher level processes of writing, with the involvement of the planning, revision and generating of ideas on paper. Furthermore, the composition of essays, poems or short stories produces writing material which varies in terms of the ideas and sentence structure among the participants, making the assessment of these difficult because of the higher level processes involved. Moreover, the higher level processes of writing hide the lower level cognitive processes of writing, which is what this study is interested in examining. To be able to investigate the cognitive processes and to compare participants, the writing tasks should be able to provide the same instructions and stimuli. This means that the writing output produced would be consistent across all participants, making it easier to compare the results. Hence the reason why this study chose copying: it provides a consistency in terms of

the stimuli provided to the participants, and in terms of the output produced. The only differences that may occur with this writing task, and which allow for easier comparison, are in the pause lengths produced and the different chunking sizes (i.e. the position where pauses occur), which are unique to an individual's language knowledge.

Copying is a simple and efficient writing task to use for measuring the cognitive processes because it gives consistent data. Using copying provides this study with a clear target; it can be known precisely ahead of time, what words or letters the participants are meant to be writing, which allows for the results of the copying to be easily analysed. Copying is clearly different to essay composition because the evaluation and analysis of essays is less straightforward, as the essay composition involves the generation of ideas, and different people have different ideas and different styles of writing. Hence, marking essays is a challenge for both human and machine judges (e.g. e-rater). Moreover, in copying, the same stimuli are used for the investigation, making it consistent for the evaluation and analysis stage. On top of that, the complexity of the stimulus can be varied so that the participants would be encountering words that they are not familiar with but would be able to naturally produce.

One might expect that copying would actually reveal very little that would be of interest to investigating the nature of processing; because all of the information is available in front of the participants, one might be sceptical as to whether it could assess their knowledge. Further, the participant does not have to fill in a gap, make a choice or even construct a sentence. Moreover, with copying, all participants will be producing the same output, so how is it possible to measure their components of cognitive processes? Research, however, has shown that the competence of the participants is still apparent even in a task like copying (e.g. Cheng & Rojas-Anaya, 2007), and this is made possible with the availability of tools such as *pause analysis*. The nature of processing is apparent through the duration of pauses between chunks, where a chunk may be a letter, a syllable, a word or even a whole phrase, depending on the level of expertise. A similar copying approach was also found useful and implemented in the studies of Cheng and Rojas-Anaya (2006 & 2007), van Genuchten et al. (2009), Obaidallah & Cheng (2009) and van Genuchten & Cheng (2010), who all explored pauses in relation to copying, words and chunking strategies.

In this study, participants were asked to copy printed sentences, visible to them one at a time, onto a piece of paper. Copying sentences allows the capture of processes such as the cognitive processes involving comprehension and word recognition that is determined by the size of

chunking (i.e. letter, syllable, word or phrase) as well as the pause lengths produced between each individual chunk. It is predicted that these processes, represented by the different pause lengths, will be able to provide information about the components of cognitive processes in language processing. This research therefore needs to establish what kind of tasks would be suitable in terms of the nature of sentence-copying. Further literature on the nature of copying will be reviewed in the next chapter (Chapter 2).

Different methods of copying sentences were explored, as well as how the copying activity can be controlled to make sure that a consistent approach is taken across all stimuli. For these purposes, this study compares *Immediate Copying* – wherein participants are asked to immediately copy the sentences when the test instructor says ‘start’, and *Initial Reading*, wherein the participants are allowed up to 30 seconds to read the stimuli in advance, before copying them. Both methods are explored in the pilot (Chapter 4), and the method that produces the most effective results is then used again in Experiments 1, 2 and 3 (Chapters 5, 6 and 7) to confirm its reliability.

## 1.2 What Kinds of Stimuli are Used?

Next, this research asks what is copied. The items to be copied should represent a sufficient amount of knowledge of the language, for example, the correct use of grammar or the ability to allow one to comprehend, in order to allow the investigation of cognitive processes in language processing. The copying materials could be a whole paragraph, a sentence, or a list of words. Copying a whole paragraph might not be suitable as it may be too long for this purpose of exploration, and may initiate issues of fatigue. Copying a list of words, on the other hand, may be assessing word knowledge only, and might be too focused as an investigation on the cognitive processes of language processing. Furthermore, the copying of sentences involves a higher level processing of meaning than the copying of arbitrary words. Therefore, this study chose a *sentence* as it contains an adequate amount of language knowledge, enabling participants to be assessed by their comprehension level at sentence level and by their familiarity with the words and grammar structure used within that sentence. There is no need for sentences to be unnecessarily sophisticated, so the sentences used for this study contained 15 to 20 words, which is considered to be adequate in order to get valuable information on the participants’ competence levels.

This leads to further questions: What kind of sentence is suitable for the copying task? Since this study takes a relatively novel approach, there are no indications from previous research as to what

kind of sentences would be suitable. The pilot and Experiment 1 therefore explore the different kinds of sentences (e.g. technical sentences, proverbs, garden path sentences) that might allow different natures of processing to be investigated. Parts of the sentences are manipulated so as to alter the difficulty of comprehending the sentence or recognising words. Furthermore, the level of unfamiliarity or effort in copying these sentences is manipulated through the omission of spaces between words, and the use of low frequency words, high frequency bigrams and proverbs. Following this initial sentence exploration, Chapter 6 is focused more on sentence types, and controls the sentence design by considering only the jumbling conditions. In the sentence design, the manipulations involve the jumbling of grammatical structure and word recognition. The use of jumbling manipulations suggests that the sentence stimuli could be assessed based on the level of familiarity of the sentence or words used, which would then affect how people with different levels of competence copy these sentences, eventually affecting the production of different pause lengths.

### 1.3 What Constitutes a Pause?

Schilperoord (1996) reported that it is not easy to define what constitutes a pause and the position of the pauses, although this is reflected by his work on *text production*. The difficulty in defining these resulted in him exploring pauses and making many assumptions with regard to their occurrences, which then motivated other researchers to explore the same. However, the literature, such as the work that involves keystroke logging (e.g. Spelman Miller, 2000, 2006; Wengelin, 2006) and graphics tablets (e.g. Cheng & Rojas-Anaya, 2006, 2007) is seen to define a pause depending on the tools used to capture the particular pauses. Certainly, defining a pause using a keyboard would mean *the time between the key-pressing of a keyboard button to the lifting of that particular keyboard button*. In the case of this study, since graphics tablets are used to capture the activities of handwriting, the *pause* is defined by the researcher as:

*‘the duration of the time between the lifting of the pen at the end of the current stroke and the placing of the pen to begin the following stroke.’*

In order to use pauses as a way to investigate the cognitive processes of language processing, it is important to know the low level components of language that could be represented by a pause. These components, as illustrated in Figure 1.1, are used as a basis to distinguish the pauses. In the process of analysing the data, it is not possible to identify the exact position of where an individual

would perform the chunking processes, as the chunking decision would depend greatly on the levels of competence, hence the reason why this study distinguishes the pauses based on the pause levels (refer to Figure 1.1). From the pause levels, the analysis will compare the pause lengths produced by every participant. The pause lengths help to provide information of whether one is taking a longer or shorter time to process the copying at that particular pause level position. It may be that long L2 pauses are occurring at word level because the participants are familiar with the words and are able to capture the whole word, or it may be that L2 pauses are long because it is the beginning of a syllable chunking that coincidentally occurs at the beginning of a word, produced by a less competent individual. L1 pauses which occur within a word may be long or short, with long L1s indicating the chunking of a syllable and short L1s indicating the copying of individual letters. The uncertainties of where particular chunking may occur for each participant at the pause level position, emphasizes the reason why it is not possible to predict definite pause positioning for all participants. Categorising the low level components of language based on the pause levels (i.e. L0, L1 and L2) may therefore be the best solution for this pause analysis.

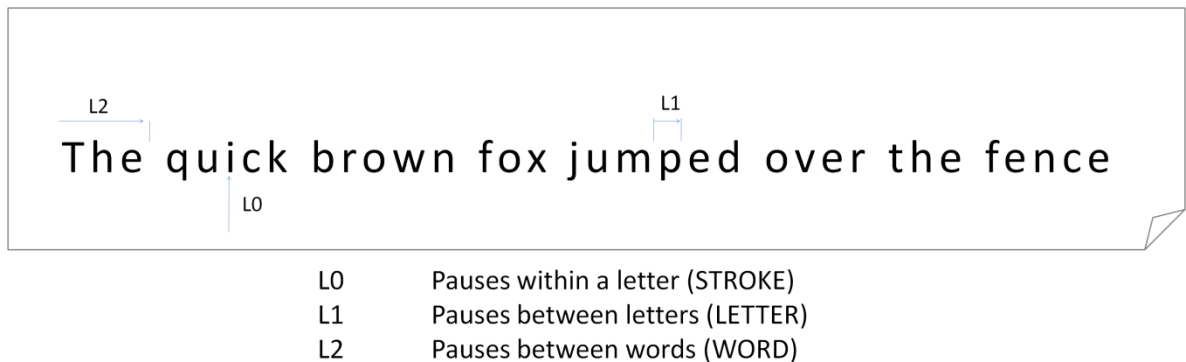


Figure 1.1 Illustration of the pause levels used in this study

The pause analysis method used in this study allows the capturing of pauses between strokes within a letter, and those between letters, syllables, words, phrases and even sentences, which in total provides richer raw data than offline, tests. These strokes that make up a letter, a word, a phrase and a sentence are a form of low level cognitive processes of writing. The term *syllable* used here in this thesis represents speech sounds either in English or in any other language used as the participant's first language. A syllable can be defined as a group of letters that makes a speech sound, which can be meaningful or non-meaningful. For some, a syllable is made (i.e. a number of letters are grouped together) based on where letters appear together in their first

language, but for others a syllable may be formed because of the ease of producing the speech sound (Zhang & Yin, 2009).

Figure 1.1 shows the lower level components that this research is interested in: stroke, letter, word and phrase. Each of these components is represented as an L0, L1, L2 and L3 pause, respectively. As shown by the arrow, the pause values would normally increase as the size of the components increases. Not shown in the figure are pauses at syllable level and pauses at word group level. This is because there is no way of knowing where these pauses may occur before the activity of copying starts, as different people have different levels of ability in the language being tested. This affects how many words or letters they can read and remember at one time before they are able to copy them. Some relate this ability to the term 'chunk': it is claimed that the working memory can hold four 'chunks' of information (plus or minus two) at one time (Cowan, 2000). Cowan (2000) defines a chunk as *'a collection of concepts that have strong associations to one another and much weaker associations to other chunks concurrently in use'*. The number of chunks that someone can remember depends on their competence level. These chunks, however, may vary in size. For example, those with good exposure to English may be able to group a number of words together as one big chunk, but those with little exposure may only be able to do so at letter level, with each chunk being a number of letters together (i.e. syllabic chunking) or a single letter.

In this research, syllable level pauses are represented as an addition to L1 pauses, shown between groups of letters. For example, the word 'jumped' may be chunked into two groups of syllables, by someone who is not familiar with the language: 'jum' + 'ped'. This may be the case for those for whom English is their second language; they might try to relate the spelling of the word 'jumped' to words in their first language or words that have the same speech sound. For example, someone with Malay as their first language might spell out 'jumped' as 'jum' (which means 'lets' in Malay) and 'ped' (which has the same speech sound as 'pad'). This method of relating the unfamiliar word to a word or sound that they know in their first language assists their copying process. Although it may seem like smooth copying, it actually shows their incompetence in the knowledge of English common syllables. In fact, the inability to recognise the word 'jump' as the root word already shows the individual's lack of language competence. Giving the same word to someone who knows a little bit of English would enable them to chunk or group the letters as the root word 'jump' and the past tense 'ed' at the end.

Word group level pauses, on the other hand, are represented as an addition to L2, shown between groups of words (e.g. the quick brown fox + jumped over the fence).

These pauses are captured in milliseconds (*ms*), and are the values that are used for the investigation of cognitive processes. The definition of the pause and what it can do compels this study to explore whether they can be used to study the components of cognitive processes at a low-level perspective, during the process of copying. The idea is supported further by Schilperoord (1996, p87), who states that '*pause analysis is a powerful tool to study cognitive aspects of language production*'. This research therefore proposes examining the pauses produced from copying sentences – using *pause analysis* to investigate the components of cognitive processes in language processing.

#### 1.4 What are the Instruments Used to Measure Pauses?

Pause analysis is the method used in this study to calculate pauses during the activity of sentence-copying. The reason for choosing pause analysis is that it offers the capturing of handwriting activity and allows the analysis of the lower level cognitive processes of writing. It offers insights into the underlying cognitive processes that take place during the activity of sentence-copying, thus offering the opportunity to investigate the nature of language processing. Pause analysis produces unique data representing each individual participant, because of the different chunking sizes and different pause lengths produced. The pauses are unique because individuals have different levels of language competence which drives how they perform the chunking.

A standard graphics tablet was used to capture the pauses from the activity of writing (Wacom, Intuous3). This was connected to a computer, and a program called the TRACE (Cheng & Rojas-Anaya, 2004). This is an *online* method, as it is able to capture results during the process of writing. The writing stimulus provided for the copying activity was a paper printed with a grid of 20 x 13 cells or boxes, each measuring 1cm x 1cm.



Figure 1.2 The writing stimulus

Participants were asked to write each individual letter in a separate box, and to ignore any spaces between words, rather than skipping boxes for this. These boxes assist in the analysis stage by ensuring that the writing of each individual letter is clear and can be easily distinguished. Once the copying activities were completed, all the data were prepared using TRACE. Correlations were calculated using Microsoft Excel. Figure 1.3 shows an example of the captured writing, in which red and pink circles indicate the beginning and end of strokes. The pauses that this thesis refers to are the grey lines between each stroke.

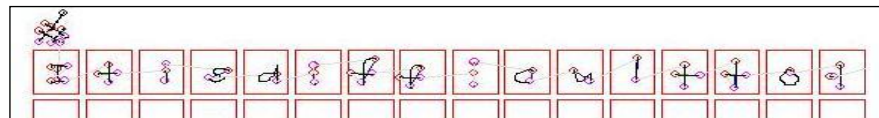


Figure 1.3 An example of the written stimuli captured and extracted in TRACE

This approach of capturing pauses using a graphics tablet and enabling the recording and analysing of information about cognitive processes of writing is called Graphical Protocol Analysis (GPA) (Cheng & Rojas-Anaya, 2004). The benefit of GPA lies in its ability to capture writing in a more natural way, because of the involvement of handwriting, than typing. GPA not only offers the potential of examining the lower-level cognitive processes of writing, but also provides other potential benefits that include: the use of modern and economical graphics tablet technology; the capture of rich raw data, with accuracy and precision; the automatic initial extraction, analysis and coding of digital behaviour protocols by computer (the relevant tools are research prototypes); the ability to capture multiple chunks simultaneously; and, finally, the use of naturalistic tasks in an experimental context (Cheng & Rojas-Anaya, 2008). In terms of pause analysis, the pauses are



captured at various levels of detail: between marks or strokes in a letter, between letters in a word and between words in a sentence.

### **1.5 How Can Pauses be Used to Investigate the Components of Cognitive Processes in Language Processing of ESLs?**

Equally important to this exploration is how the pauses are measured. With the capturing of pauses at every stroke of writing and the categorisation of these pauses into different pause levels, one can imagine how huge the amount of raw pause data produced is. So, how can this study calculate these pauses? What statistical measure would be suitable in performing the calculation? With the ability to capture and distinguish the pauses according to the pause levels (i.e. L0, L1, L2), it allows the measure to be made based on the pause distributions, suggesting the possibility of applying measures such as the median, mean or quartiles in search of the suitable measure for cognitive processes. Additionally, in measuring pauses, this research must allow for individual differences, in that people write at different speeds. This may well affect any measure that this research finds. One way of dealing with this may be to calculate individual baselines, by providing a test to capture the basic writing speed of each individual participant, in which there would be minimal language processing. This could be used to normalise individual data, which allows the compensation for different writing speeds. All these measures will be explored through a series of experiments.

The sentence-copying activity allows the study to investigate the components of cognitive processes in language processing based on the chunk sizes (represented by the pause values), which are identified and categorised according to the different elements of the sentence (i.e. letter, syllable, word and phrase). These chunks correspond to the pauses, which may be of different lengths and may provide information about the nature of language processing. As mentioned, the statistical measures explored may reveal important information from the raw data captured. This is on the basis that there are many different levels of low level components of language, such as the letters and words, which are distributed across a sentence, so a measure must be found that could be an effective summary of the pause distributions that distinguish people with different natures of processing. Pause distributions are skewed: longer pauses are associated with breaks between chunks, before a chunk is copied, and shorter pauses occur within a chunk.

Somebody who is very competent and has a greater familiarity level with the sentence and word characteristics could be expected to be able to process a larger chunk size and produce a greater number of long pauses, while somebody who is less competent and less familiar with the sentence and word characteristics might process a smaller chunk size and produce greater numbers of long pauses.

An effective measure should be able to clearly differentiate the nature of language processing within the participants recruited. One way of doing this is by correlating the measure of pauses with a score of an independent language test. This approach of validating the method is common to Language Testing researchers. Performing the validation helps to establish whether sentence-copying is a reliable approach. The independent language tests on which this study are based are the Oxford Diagnostic Test (ODT) (introduced in Experiment 1) and the Vocabulary Size Test (VST) (introduced in Experiment 2). Each of these assesses different aspects of language: grammar knowledge (ODT) and vocabulary knowledge (VST). These aspects are referred to in this research as the *components* of language competence. These *components* are considered critical to language competence. Although common to many, the correlations obtained from the two components may not be the same value, because they measure different aspects of the language. This approach to validating the sentence-copying will also assist in confirming whether the method is able to investigate the language components similar to those being assessed in ODT and VST. Over the course of the various experiments, this study will therefore look at general competence (pilot study), competence with respect to grammatical knowledge (Experiments 1, 2 & 3), and competence with respect to vocabulary knowledge (Experiments 2 & 3). The natures of language processing are measured through the different pause lengths produced from the alternative sentences which have been variously manipulated to change their levels of difficulty.

## 1.6 How do you Validate the Measure?

An important question that must be asked is: ***“Does the pause measure what it is intended to measure?”***(Lado 1965: 30). In order to answer this, test validation is required. In language testing, test validation refers to being able to establish a reasonable link between a test-taker’s performance and their actual language ability (Kluitmann, 2008). There are many ways to validate a test and there is no doubt that implementing as many validation methods as possible would ensure a better result (Alderson et al., 2005). But what interests this thesis in relation to the

experiments conducted is the *external validity* of a test, where '***a comparison of the test scores with some other measure for the same test taker taken at roughly the same time as the test***' is made (Alderson et al., 1995). This *other measure*, against which the data is correlated, could be a parallel version of the same test, some other test that measures other components of language, a performance ranking order given by teachers based on language ability, or a ranking order made based on biographic language background. The important point is that such validation is required to ensure that the test is measuring what it is intended to measure.

The most frequently-used statistical measure is the *correlation coefficient*. The correlation value is able to show whether two tests are assessing similar or different aspects of language knowledge. The most common type of validation is the *concurrent validation*, usually used when there is a need to develop a test to replace similar tests that are expensive or difficult to obtain. Concurrent validity coefficients range from +0.5 to +0.7 (Alderson et al., 1995). A correlation of +0.9 would indicate that a strongly related aspect is being tested.

Another aspect of validation that is heavily associated with this study is the *internal correlation* that belongs to the *construct validity*. Internal correlations focus on relating the different test components (or sections) with each other (Alderson et al., 2005: 184). The aim is that all test components should measure something different, which would then contribute to the overall picture of language ability. Hence, the correlation expected should be fairly low, +0.3 to +0.5. If the correlations between test components are found to be high, at +0.9, this would mean that the components are testing similar things, thus one of the two should be dropped. The validation would also establish whether the test is assessing similar or different attributes (Alderson et al., 2005). The correlation between each sub-component and the whole test, however, should normally be at +0.7.

Data from several research studies has identified similar correlation values to those given above. For example, Salah (2008) reported a correlation of +0.7 and +0.6 between the percentage of known words and students' comprehension of two reading tests. Koda (1989) and Khailidieh (2001) both reported strong correlations between vocabulary tests and reading comprehension in the Japanese and Arabic language. Studies that involve word knowledge and reading comprehension are likely to fall between the correlations of +0.3 to +0.8 (Tannenbaum et al., 2006).

To conclude, in order to confirm whether a test is actually measuring what it is intended to, a test validation will need to be performed. Even if the test developer is unsure of the exact skills being tested, the running of a ‘construct validity’ test would assist in **‘empirically testing hypothesized relationships between test scores and abilities’** (Bachman, 1990: 256). Further, the studies found in the literature, as described above, assist in setting the criteria to validate the novel method developed in this study, which compares pauses during sentence-copying to the independent English tests.

## 1.7 Why focus on Second Language Learners?

The reason behind the choice of participants is with respect to this assumption:

*“Someone who is very competent and has a greater familiarity level with the sentence and word characteristics could be expected to be able to process a larger chunk size and produce a greater number of long pauses, whilst someone who is less competent and less familiar with the sentence and word characteristics might process a smaller chunk size and produce greater numbers of long pauses.”*

The main focus in recruiting participants is to have a range of different competences to provide a contrast for the experiments, but the range need not vary much in other respects, such as in the case of mixing First Language Learners (FL) and Second Language Learners (SL). If you use adult native speakers from a wide range of educational levels, this may introduce different levels of intelligence as a confounding variable. Using children and adults is worse, because children have less practice in writing.

Thus, educated adult degree students with ESL are a good population, because they are in most respects similar other than language experience. Factors that may differ amongst participants would most likely be connected to their level of exposure to English: how familiar they are with the words and how well they comprehend what is being copied. As a whole, this would then be able to assist the investigation of components of cognitive processes in language processing and determining participant’s nature of processing. The next chapter will consider the nature of SL processes as they relate to the components affecting the processing of copying.

## 1.8 An Example of How the Cognitive Processes in Language Processing in ESLs is Measured

As a light introduction to the work in this study, a sample sentence is presented, which participants would be asked to copy as quickly as possible. It raises a number of interesting questions.

*'thetheoristintimidatedistraughthonourablerevuedramatisenarrator.'*

The sample sentence given above has been manipulated in such a way that the identification of the words is made more difficult, through the use of high frequency bigrams at the beginning and end letters of every word, and the omission of spaces between words. This then nurtures questions such as:

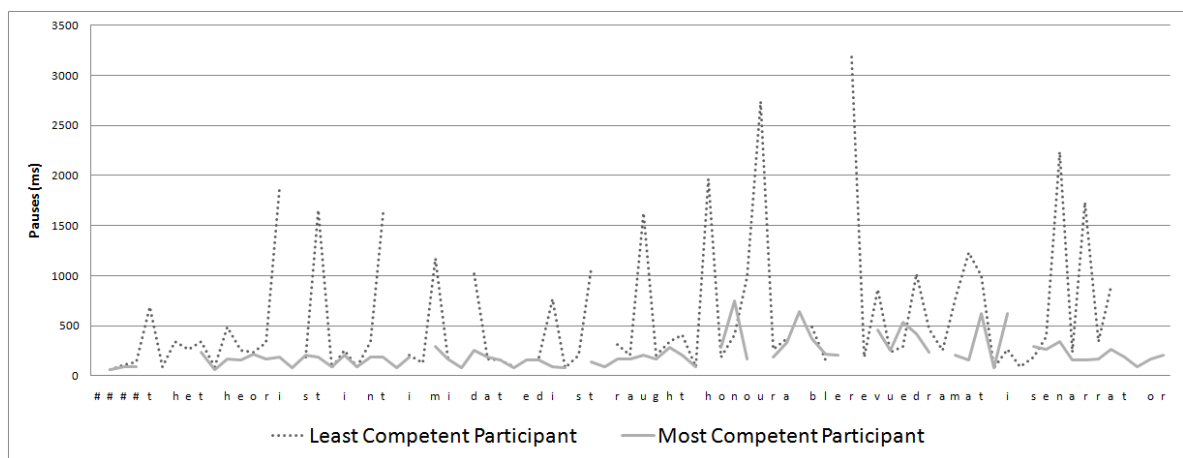
- *In the process of copying*, where would the chunking occur? Would it occur between words? If so, with the condition given above, how easy is it to distinguish the words?
- Would there be any difference between someone who is competent in the language and someone who is less competent?
- Would a competent person still be able to capture a group of words in one chunk, or would they be confused by the bigrams used at the beginning and end letters of every word?
- Would a less competent participant be able to distinguish the words? If not, what kind of chunking would they adopt? Would they copy letter-by-letter, or by groups of letters together, i.e. syllables?

These questions are all connected to the chunking processes strategy: how does the process of chunking affect the pauses?

Firstly, processes such as the decision whether to chunk at word level or letter level would take some time. Although measured in *ms*, it is able to show differences between participants which are unique to every individual depending on their level of competence and the nature of language processing. A competent individual would produce short pause lengths, as they are familiar with the language, and a less competent individual would produce longer pause lengths because of less familiarity with the language. However, a sentence manipulated such as the above would definitely increase the pause lengths for all levels of competence. Although this is the case, this

study might still be able to differentiate between the levels of competence, because a competent individual might be processing the copying meaningfully, hence the chunking may occur at word level, but a less competent person, who finds it difficult to distinguish the words, might be producing many small chunks at syllable level, producing many long pauses between syllables.

These are among the interesting questions raised which will be explored through a course of experiments. As a general idea of the sample data gained from the copying of the sentence above, an illustration is given below, Figure 1.4, which provides the overall pause distributions across the successive letters in the sentence.



**Figure 1.4 An example of how pauses differ between the most and least competent within a group of participants**

In Figure 1.4, pauses are presented on the y-axis, in milliseconds, and the successive letters are on the x-axis. The solid line represents the Most Competent (MC) participant and the dotted line represents the Least Competent (LC), from the group of participants recruited. As can be seen, LC has many long peaks of above 500ms. These long pauses of LC occur at L2 (word) level and L1 (letter or syllable) level. MC, on the other hand, has a smaller range of pauses, mostly below 500ms for all levels of pauses; L2 (word), L1 (letter or syllable) and L0 (stroke) level.

As a general example to explain the position and occurrence of pauses, the word ‘the’ (as can be seen with LC) has a longer pause at ‘t’ and shorter pauses thereafter for ‘h’ and ‘e’, showing a clear L2 (word) level pause. To observe the pauses within a letter (L0), take for example the letter ‘t’. A long pause occurs before writing the first stroke of ‘t’, and the next pause value suddenly drops because the second stroke of the letter ‘t’ is an automatic process. For an illustration of a syllable chunking, in the word ‘honourable’ (in the case of LC), it is observed that a long pause occurs at ‘h’

and shorter pauses at 'ono'. The pauses increase again from 'u' until the writing of the word is completed. Two parts of syllable chunking of the word 'honourable' were seen: 'hono' and 'urable'.

The overall graph shows a clear distinction between MC and LC in processing the copying of this kind of sentence. The long pauses reveal the difficulty in processing the individual chunks, for example the difficulty in distinguishing the words because of the space omission and the use of bigrams. As shown, the examples were able to distinguish the two levels of competence by examining both the different lengths of pauses and the positions where chunking occurred. This is a strong indicator that pause analysis is a promising approach to measuring the cognitive processes of language processing. Taking this into account, the researcher will focus on the analysis of pauses, and examine how this study can measure the cognitive processes using these pauses in writing.

## 1.9 What are the Underlying Cognitive Processes Involved?

As measuring the components of cognitive processes in language processing using sentence-copying is a complicated process, there is a need to understand the different processes employed during the copying activity. This leads this study to the development of a model: the theoretical Model of Copying (MoC). The model not only helps in understanding and explaining the occurrences of copying activities and the pauses produced, but also allows sentence design by making predictions, to see if the sentences are suitable.

The main challenge in this study is that of understanding what generates the different pause lengths for different natures of language processing, so a model that could systematically explain the processes involved during copying was developed (c.f. Chapter 4). This model was developed after conducting the pilot and Experiment 1. As well as being used to interpret results, the model was used in the process of designing the stimuli. Sentences were simulated and, through the model, predictions were made as to the possibilities of chunk sizes, and these predictions were analysed. The aim of the simulation is to examine if the sentences are able to differentiate the natures of language processing. Through the simulations, the different patterns of processing a participant might take when copying could also be identified.

From the model, it is also assumed that the length of a pause is created from the occurrence of cognitive processes, but that these processes may occur either consecutively or simultaneously. It is difficult to define the individual lengths of a specific process, unless a thorough study is made specifically in defining or validating the model alone. This may provide scope for future research which takes into consideration the findings and data of this study, involving the L0, L1 and L2 pauses.

## 1.10 Conclusion

With this in mind, this study therefore chose to focus on copying sentences within the context of free handwriting, and focus only on adults with English as their second language (ESL). The factors discussed above will be explored through a series of empirical studies (Chapters 4, 5, 6 and 7). It is empirical because the research is based on observed and measured phenomena occurring in the series of experiments. The results of each experiment generate questions for the next experiment. In the case of this thesis, such research work is novel, hence various possible approaches are explored, as in the first initial exploration in the pilot study, in order to test if it is feasible to use pause analysis to measure the cognitive processes. The approach is then to tackle interesting issues that are of strong interest to the objective of this research. During the exploration, a Theoretical Model of Copying (i.e. MoC) was developed as a means to understand the processes that occur during copying. This not only assists understanding but also helps in terms of choosing suitable sentences for copying in the subsequent experiments.



Therefore, the overall aim of this study is:

**TO DEVELOP A NOVEL METHOD FOR ASSESSING COMPONENTS OF COGNITIVE  
PROCESSES IN LANGUAGE PROCESSING BY ANALYSING THE PAUSES THAT  
OCCURS DURING THE ACTIVITY OF SENTENCE-COPYING.**

In conclusion, this thesis explores several variables, including:

- a) the method of copying: whether some Initial Reading (IR) time is given for participants to prepare themselves before the actual copying begins or whether copying begins immediately, without any initial preparation, Immediate Copying (IC).
- b) the kinds of sentences to be copied: whether the sentences are ordinary everyday sentences, proverbs, technical sentences or garden path sentences, for example.
- c) the possible measures of cognitive processes: the different levels of pauses that determine the processes related to language and the different statistical measures that may serve as particular measures.

To summarise, this study hopes to achieve the overall aim of this research work: *to develop a novel method for assessing components of cognitive processes in language processing by analysing the pauses that occur during the activity of sentence-copying.*

## **CHAPTER 2    Literature Review: Studies of Pauses, Sentence-Copying & Second Language Processing**

### **2.1 Introduction**

This chapter presents the current literature involving the three major research areas which have contributed to the development of the novel approach: (1) studies of pauses, (2) sentence-copying and (3) second language processing.

Section 2.2 will focus on the potential of pause analysis as a tool for measuring cognitive processes. Literature is examined concerning the current development and benefits of pause analysis, which has produced a number of applications. Research in pause studies has involved tasks such as composition and dictation. This thesis, however, will focus only on the 'copying' technique, which assists in revealing the lower-level cognitive processes of writing.

Therefore, section 2.3 will discuss the literature in relation to copying as a technique; the processes of copying which involve reading and writing; and how copying can contribute to the measuring of cognitive processes.

Section 2.4 focuses on second language processing. This section will review aspects that might influence how the participants process their second language. In order to do this, there is a need to study factors that might affect the way one processes the reading of the sentences before the copying begins.

Finally, in order to identify what contributes to the length of pauses and how these can reveal the nature of processing, Section 2.5 will discuss the factors that can affect the cognitive demands of copying.

### **2.2 Pause Analysis: An Alternative Method**

Studies on writing have been conducted since the 1970s and range from analysing products (e.g. written documents) to using real-time methodologies (Schilperoord, 2001). Researchers quickly realised that it was impossible to understand what takes place in the writer's mind during composition by just observing the texts and documents already produced. Hence, writing research began to shift towards the analysis of the underlying cognitive processes that occur while writing.

This is vigorously explored by a number of researchers, with the discovery of tools, methods and software that assist in capturing pauses. These developments are discussed further below.

Meanwhile, test developers in the Language Testing field struggled to find ways to understand and capture the processes that take place during test-taking. These researches proposes that pause analysis and sentence-copying are valid alternatives to the former investigation of the cognitive processes of language processing approaches, and a step into enabling participants to be assessed online, during the copying activities. The sentence-copying tasks allow the capturing of cognitive processes, such as the semantic processing for retrieving meanings, compared to the usual test items, which measure knowledge offline and produce test scores, i.e. ‘products’.

This section will therefore look briefly into how the study of pause analysis was discovered, examine the role of computers and technology, and finally assess how pauses can reflect cognitive processes.

### 2.2.1 The Initial Approach

In the 1980s, Hayes and Flower were among the first to initiate the experimental process-oriented approach to testing writing, which revolutionised the traditional product-oriented based studies (Schilperoord, 2001). The primary method to analyse the procedure a writer engaged in while composing was the use of case studies (Schilperoord, 2001). Matsushashi (1981, 1982, 1987) proposed the ‘think-aloud protocols’ and the videograph technique as alternatives to the observation of writing. Instead of conducting analysis solely based on transcribing speeches, the videograph technique allows the observation of the writer’s behaviour during writing, such as their hand movement or facial expressions. Writers were witnessed as they *‘propell[ed] the text forward through space and time’* (Matsushashi, 1982: 270). The videograph required a writer to describe out loud the thinking processes s/he underwent while composing. This was the first major technique available that acted as a ‘window’ to the cognitive processes underlying text production (Schilperoord, 2001). It was, however, criticised in terms of reliability and validity (e.g. Cooper & Holzman, 1983; Kowal & O’Connell, 1987; Russo, Johnson & Stephens, 1989).

Building on the breakthrough of Hayes and Flower, researchers (e.g. Schilperoord, 1996; Spelman Miller, 2000a & 2000b) developed methods for capturing writing activities in real-time. This has been referred to by researchers (e.g. Spelman Miller, 2000a, 2000b, 2006; Cheng & Rojas-Anaya, 2004, 2005, 2006, 2007, 2008; Spelman Miller & Sullivan, 2006) as the ‘temporal course’ or ‘pause’

of the text production (the latter term will be used commonly in this thesis). Defining a pause depends highly on the methods implemented, as discussed in Section 1.2.1. Research has also been able to prove that pauses reflect cognitive processes, as explained in Section 2.3.3. The pause analysis technique was influenced and driven by studies in speech production in the 1960s. This was a long researched field, established before pause analysis was used in writing (Schilperoord, 1996; Spelman Miller, 2000a, 2000b, 2006). The work of Schilperoord (2002) sampled a pause database, which underpinned two sampling methods: the auditory recording of pauses and the keystroke recording methods. His findings on the multiple types of pauses in speech, such as hesitation phenomena, restarts, self-corrections, and filled and unfilled pauses, has motivated further work in defining the position of pauses in text production (Spelman Miller, 2006). Even though these works are not closely related to pause analysis in terms of the lower level cognitive processes of writing, they provide a foundation for analysis. In other words, the starting point for pause analysis lay in the effort to uncover the cognitive processes underlying text production. The great potential of this approach remains untapped, however, which is why this thesis attempts to explore it as a tool to measure the components of cognitive processes of language processing.

### 2.2.2 The Role of Computers & Technology

The last decade of the 20<sup>th</sup> century saw major developments in Studies of Writing research due to the growing popularity of computers. A new method for recording writing research, which allows researchers to investigate the cognitive aspects of writing online, was introduced; keystroke logging (Spelman Miller, 2000; Schilperoord, 2001 and Wengelin, 2002). It is not clear exactly when keystroke logging was introduced, but some initial work was undertaken in the mid-1980s by Bridwell, Sirc and Brooke (1985) and Bridwell-Bowles, Johnson and Brehe (1987), who used the 'Recording WordStar' to study word-processor writing at the University of Minnesota (Spelman Miller and Sullivan 2006). Since then, many other programs based on keystroke logging have been introduced: ScriptLog (Strömqvist & Malmsten, 1998, Ahlsén & Strömqvist, 1999; Strömqvist & Karlsson, 2002), Inputlog (Leitjen and Van Waes, 2005) and JEdit (for Mac users). Supported further by the advancements in computers and technology, research on the recording of the cognitive processes of writing through pause analysis can now be achieved through the use of applications such as *Graphical Protocol Analysis* (Cheng & Rojas-Anaya, 2004), as well as *Eye and*

*Pen* (Alamargot & Chesnet, 2006), which also incorporates eye tracking. Brief overviews of these applications are presented below.

### 2.2.2.1 *Keystroke Logging*

Keystroke logging refers to *'the computer recording the writing activity as writers compose on a keyboard'* (Spelman Miller and Sullivan, 2006). The theoretical underpinnings of keystroke logging research from a cognitive approach concentrate on how writers produce text that is on the higher-level writing processes involved and how far writers are aware of them at a composition level. Spelman Miller and Sullivan (2006) reported that the functions of keystroke logging involve the electronic recording of activities such as key presses, editing and cursor movements. The tool also allows for the recording of the fluency of writing, in terms of pause locations and durations and the sequence of actions during writing. The advantages of keystroke logging have been documented by many others (see van Waes, 1991; Severinson, Eklundh & Kollberg, 1996a, 1996b; Pennington, 1999; Strömquist & Ahlsen, 1999; Spelman Miller, 2000a).

In the past, the term 'writing' was strongly associated with the high-level processes of writing (i.e. planning, translating and revising). These make up what is known as composition or 'text production' in general. Writing was identified as a problem-solving activity. In the beginning, researchers in the area of keystroke logging found it difficult to differentiate the planning process and the translating and revising processes, and this was further impacted by the participants' different keyboard skills. These constraints suggest that further research is needed in the area of 'text production' (as initiated by Leijten & van Waes, 2006; Lindgren & Sullivan, 2006; Spelman Miller, 2006; Wengelin, 2006 and others). The development of keystroke logging has also introduced new ideas in respect to the study of the lower-level cognitive processes of writing. The lowest level that this method has been used for in the study of pauses, however, is at words level, in a recent study by Wengelin (2007).

To conclude, the application of keystroke logging has been very useful in many areas, such as the linguistic, textual and cognitive study of writing, the development of language learning, literacy and language pedagogy (Spelman Miller and Sullivan, 2006), the research on comparing speech and writing, the study of learning how to write, writing difficulties and the experimental study of spelling (Strömquist et al., 2006). Even though the method is closely related to Graphical Protocol Analysis (GPA), keystroke logging uses the keyboard as a medium for composing, whereas GPA

allows the natural handwriting to be recorded via a graphics tablet. The use of handwriting is common in schools, is considered unique in representing a particular individual and is a basic skill for everyone who is educated. The pauses associated with it are therefore, more distinctive. Keystroke logging, on the other hand, would require an individual to learn the keyboard skills before being able to use it.

#### 2.2.2.2 *Graphical Protocol Analysis (GPA)*

The term *Graphical Protocol Analysis* (GPA) refers to a methodology that focuses on the duration of pauses and the analysis of the time frame. It defines pauses as ‘the duration of the time between the placing of the pen to begin the current stroke and the lifting of the pen at the end of the following stroke’. The initial aim of GPA is slightly different to that of keystroke logging, in that it exploits the measurement of pauses to study whether the chunking of information in memory can be revealed by pauses in writing and drawing (i.e. graphical production). During writing production, pauses can be captured at various levels: within a letter (L0), between letters (L1), between words (L2), between phrases (L3) or between sentences (L4). It has been claimed that pauses in graphical production reflect the hierarchical organisation of chunks in working memory and thereby produce a *temporal chunk signal* (TCS) (i.e. the pauses) that can provide information about the organisation of the chunks (Cheng & Rojas-Anaya, 2008; van Genuchten & Cheng, 2010). In the studies of Cheng, Rojas-Anaya and van Genuchten, the magnitude of pauses is significantly and substantially greater for higher levels in the chunk hierarchy. For instance, with language-like stimuli the duration of pauses in one sentence increased successively over the stroke, letter, word, and phrase levels. Based on these findings, it has been considered possible to apply pause analysis to measuring the cognitive processes of language processing.

Among the benefits of this approach (refer to Chapter 1) over keystroke logging is the focus on free handwriting, which can offer richer raw data at a more in-depth level (pauses between strokes in a letter) (Cheng & Rojas-Anaya, 2005; Medwell & Wray, 2007). This data cannot be captured using keystrokes. Further, adults (the target participants for this study) are generally well-versed in handwriting, having learned how to write at a young age, but not everyone is competent at typing. Accordingly, a variation of typing competence might be introduced. The work by Cheng and Rojas-Anaya (2005, 2007 & 2008) is what inspired the research in this thesis, as it provides a template for using this method for the study of language assessment.

### 2.2.2.3 Eye Tracking

In a further attempt to investigate writing, the *Eye and Pen* approach was introduced by Alamargot, Chesnet, Dansac and Ros (2006). In addition to the graphics tablet, they used an eye tracker in order to capture the eye movement of participants when reading the stimuli. This not only enabled the experimenter to observe eye movements, but also assisted in tracking the time length of eye-fixation on the screen. The study of Alamargot et al. (2006) examines the activity of high-level writing, i.e. composition. The use of eye tracking assists their studies in terms of tracking the *revisions* made by the writer. A *revision* is when the participant's eye refers back to the words that they have composed, in order to assist with the continuation of the composition. A similar approach to this is the *EyeWrite* (Simpson & Torrance, 2007) combined with *ScriptLog+TimeLine* (Strömquist & Karlsson, 2002; Anderson et al., 2006), but instead of handwriting, it focuses on *keyboarding*. This approach is suitable for some writers, whose keyboarding is more fluent than their handwriting. This approach, however, has once again been used to study text production during composing, i.e. a high-level writing activity. It has not been used for measuring language competence. The benefits of EyeWrite include the synchronisation of the records of keystroke logging and eye fixation data.

A recent study by Wengelin et al. (2009) explores the combination of keystroke logging and eye tracking (adopting the EyeWrite and Scriptlog+Timeline) in order to study the cognitive processes underlying written text production. The eye tracker is used to monitor eye movement during text production at word level. The preliminary findings have shown that pausing is often accompanied by directed eye movements within the text (Wengelin et al., 2009). Wengelin et al. (2009) also discovered the *parallel processing* occurring during the writing process during the keyboarding activity through eye tracking. The eye fixates on words that have been composed, while the hand is still typing new words. The findings of parallel processing are not surprising, as there have been similar findings in speech production (Power, 1986) and written production (Chanquoy, Foulin & Fayol, 1990). Study in terms of understanding the underlying cognitive processes is still in progress (Schilperoord, 2003; Wengelin et al., 2009). Even so, these works still evolve around text production and none has yet explored sentence-copying as a tool to measure the cognitive processes.

On the other hand, a recent study by Barone (2010, work-in-progress), associated GPA with eye tracking and investigated the effects of writing under different stimulus types. The findings have

shown that there is a difference of approximately 200ms (distributed fairly across the copying activity) between the time it takes to write on a straight line and the time it takes to write each letter in an individual letter box. These findings impact the present research in that they confirm that the use of boxes to write individual letters will not confound the results. In brief, it can be summarised that the exploration of eye tracking contributes to the understanding of the processes that underlie written composition. Such an approach, however, would be too wide for this thesis, and so it will focus only on GPA. Even so, the application of eye tracking would be beneficial for future research.

#### **2.2.2.4 Conclusion: GPA in Focus**

In spite of the development of all the aforementioned tools and applications, none has yet been employed in the investigation of components of cognitive processes in language processing. The only exception is the recent attempt by the ETS, which employs pause analysis in examining high-level writing processes (Quinlan, 2010). This gap in the literature may be largely due to the fact that researchers are still in the phase of exploring the tools. The interest of this thesis, however, lies in the involvement of free handwriting and sentence-copying using the methodology of the GPA system, which provides rich raw data that represent the cognitive processes that might relate to language competence, due to its ability to capture online the pauses at lower levels of writing in as much detail as at L0 (stroke), L1 (letter), L2 (word) and L3 (group of words) levels. Hence, much of the work in this thesis is built upon that of Cheng & Rojas-Anaya (2005, 2006, 2007 & 2008).

#### **2.2.3 Pauses as Reflections of Cognitive Processes**

It has been well established in Cognitive Science that the duration of pauses between actions reflects the amount of mental processing that is needed to prepare for these actions (Fayol, 1998; Kellogg, 1998; Torrance & Jeffery, 1998). In his groundwork study, Schilperoord (1996) explored the distribution of pause time in text production, its relation to textual characteristics and the *cognitive structure model and processes* that underlie text production. He has further strengthened the argument that pauses reflect cognitive processes by justifying that pauses signal cognitive processes. A good understanding of the nature of pauses and the processes involved has still not been reached in research (Schilperoord, 2003). This explains the development of many tools and applications for further investigation (i.e. eye tracking and keystroke logging).



It is known that pauses during writing indicate that processes are taking place, especially when pause lengths vary. The lengths of pauses fluctuate according to the features of the text that is being composed (Schilperoord, 1996; Foulon, 1998), which indicates that pause locations are not coincidental, but instead point to mental activity associated with writing processes (Wengelin et al., 2009). Pauses involuntarily reflect cognitive processes and the variance in pause time reflects differences in cognitive processing efforts (Schilperoord, 1996). The variance in pause lengths is also reflected in the hierarchical level of processing. For instance, pauses at paragraph level are likely to be longer than pauses at sentence level, which, in turn, are longer than pauses at individual word level (Cheng & Rojas-Anaya, 2008; van Genuchten & Cheng, 2010). Schilperoord (1996) concluded further that the *'production rate parameters such as pause time differences are sensitive to the temporal course of cognitive processes in production,'* therefore, *'cognitive processes are represented in the temporal characteristics of text production'*.

In determining that pauses reflect only cognitive processes, Schilperoord (1996) disagreed with de Beaugrande (1984: 160), who stated that pauses are multi-determined phenomena (also mentioned in Butterworth, 1980 and Garman, 1990). This suggests that there is a strong need to identify other possible causes that might influence the duration of pauses. Therefore, Section 2.5 explores the possible factors that might contribute to different pause lengths, in accordance with the aims of this thesis.

Schilperoord (1996) describes pauses to be involuntary, suggesting that there is a need to make sure that no other processes are responsible for the pause. For this reason, it is crucial to design the experiments carefully, in order to control the activity of writing. Hence, this research explores sentence-copying, which allows a much focused investigation to take place, examining the lower-level cognitive processes that occur within a second. Further, sentence-copying allows the control of factors unrelated to language competence that may adversely affect the duration of pauses. It is hoped that, this way, the processes that might be involved in the processing of sentence comprehension will be understood and used to measure the cognitive processes in language processing.

### 2.3 Copying: An Insight into Understanding the Cognitive Processes

Copying was normal practice in schools in the decades preceding the development and ready availability of photocopies, scans and print-outs [c.f. Lorenz & Grabowski (2009) for a historical

perspective on the role of copying]. Copying is used both as a method to practice skills and self-control strategies (Brinkmann, 2004), and as a means of punishment and enforcement of discipline. Despite the advent of technology, copying still plays an important role as a working technique in schools (Grabowski et al., 2010). Copying has been proven to be a skill that helps improve handwriting style, motor skills, and handwriting fluency, but also word recognition, especially when it is used for practicing word spellings (Sassoon et al., 1986; Suselbeck, 2003; van Galen, 1991). Moreover, Grabowski, Blabusch and Lorenz (2007) have concluded that copying can be considered as the basis of all complex writing abilities. Therefore, copying may offer a rich insight that uncovers the underlying cognitive processes in language processing.

### 2.3.1 Why Copying?

There are indeed very few studies that focus on copying. Research on copying is normally related to the study of handwriting instruction. The earliest research was conducted in 1975 by Askov and Greff, who examined the differences between copying and tracing, in order to determine which is the most effective type of practice. Their study reveals the advantages and disadvantages of both methods, as used by children in schools. It suggests that tracing is an easy enough task, while, by contrast, copying involves more meaningful processes, which affect learning. Gonzalez et al. (2010) compared tracing and copying in the reproduction of patterns. Their findings have shown that tracing is beneficial for short-term learning and encourages the provision of accurate and immediate feedback. Copying, however, requires greater use of memory and is found to be especially useful in the long-term learning of novel letter shapes. Kirk (1980) had already arrived at the conclusion that copying is a better method for teaching children new shapes. It could be that copying has the advantage of forcing individuals into remembering the shapes (Gonzalez et al., 2010) and, when combined with the actual movements of handwriting (kinaesthetic), facilitates the visual memory of graphic shapes and letters.

Tracing does not require much effort, as writers only need to re-form letters that have already been produced. Copying letters and word forms, on the other hand, demands more effort and concentration and hence, requires more visualisation (Askov & Greff, 1975). For that reason, copying was found to be superior to tracing. As copying requires the reproduction of letters and words on a piece of paper, it requires some sort of chunking activity to take place, where writers can process bits of information at a time. Assuming that writers are well-versed in handwriting, the memory resources used in copying are most likely to represent processes in relation to the

comprehension of the copying material. Therefore, it was concluded that during sentence-copying pauses reflect meaningful processes, and so studying them could be used for measuring the components of cognitive processes. Grabowski et al. (2010) claimed further that even though copying is a low-level writing task, it involves several sub-processes that might offer rich raw data revealing the underlying cognitive processes of other areas, such as language acquisition.

Accordingly, the process of copying involves the inter-relationship between writing and reading. How do they interact? A general picture of a copying process would need someone to read first, then to copy what has been read. But a detailed scenario would be that this someone will need to chunk the reading materials into smaller chunks depending on their competence in the language. Of course, this competence varies because of different factors affecting the language user. For example, a more competent language user would be able to encode bigger chunks during reading than a less competent language user. However, will the more competent language users be able to use the big chunks in writing or will they break them down into smaller pieces? The decision also depends on the size of chunks that the working memory can hold before performing the writing process. This would indeed affect the pause length before the writing process even begins. Therefore, are reading-chunks bigger than writing-chunks? Suitably, copying may be able to provide a rich insight in order to understand this relationship further.

### **2.3.2 Previous Research in Copying**

A gap exists in the literature after the 1970s, until the early 1990s, when Rieben, Meyer and Perveaux (1991) studied copying from cognitive and instructional perspectives. They found seven strategies of copying in children, including syllables, letter writing and bigrams, among others. Even though the literature has been slow to realise the potential of linguistic features (e.g. phonemes and graphemes) in relation to copying and language processes, these features do have a role to play in the process of chunking when it comes to spelling (e.g. Rieben et al., 1991; Verhoeven et al., 2006; Kandel et al., 2009).

Research on copying was then carried further by the very recent work of Grabowski, Weinzierl and Schmitt (2010), who looked specifically into the performance of children in copying. Even though these studies involve children, there is nothing to suggest that the processes and strategies observed are not adopted by adults. The study by Grabowski et al. (2010) found that performance

in copying improves as children get older, especially when recognition and handwriting become automatic.

As well as the application of the copying technique as a method to improve one's skills, it has been used in a number of studies, such as understanding the relationship between early reading and writing skills by investigating the copying strategies of children (Rieben et al., 1991; Saada-Robert & Rieben, 1993; Rieben & Saada-Robert, 1997), measuring the fluency of alphabet writing in schools in order to assess handwriting competence in children (e.g. Longcamp et al., 2003, 2005, 2006, 2008; Rosenblum, 2005), investigating chunking strategies at stroke, letter and word level (Cheng & Rojas-Anaya 2005, 2006), exploring the copying unit size of children (i.e. phonology and orthography effects) from a French elementary school (Kandel & Valdois, 2006a) and then comparing their findings with children from a Spanish school (Kandel & Valdois, 2006b), investigating the effects of syllables in the process of segmenting words during copying (Verhoeven et al., 2006), measuring the competence of four different participants in writing mathematical formulae (Cheng & Rojas-Anaya, 2008), and examining typing skills proficiency in adults (Grabowski, 2008). The most recent research is that by Grabowski et al. (2010) as described above. So far, there has been little discussion on the application of copying as a tool to investigate the components of cognitive processes in language processing.

### **2.3.3 The Process of Copying & Working Memory**

The process of copying in this research refers to the involvement of working memory capacity, handwriting, and word and meaning related processes; which will be discussed in Section 2.5. How successful one is in copying fluently depends on various reasons, such as those explained in Section 2.4. The activity of copying itself demands resources from the working memory, but how much depends on the condition (i.e. language competence) of the participants. The execution of a copying task is expected to be similar to the activity of high level writing (e.g. essays) in terms of its dependent on the availability of working memory capacity (e.g. Kellogg, 1996, 1999), except that copying is a reduction of writing to low-level processes; from lexical retrieval processes (i.e. word recognition) to execution processes (i.e. the activity of writing the words or copying) as claimed by Grabowski, 2008.

The working memory capacity is limited (e.g. Baddeley & Logie, 1999). The availability of the capacity depends on how a person processes information. For example, someone who is fluent in

handwritten activities would not need the space in working memory to remember how to form each letter; they could produce the letters automatically. However, someone who is unfamiliar with the letter form would require extra space to remember the shape of the letter, which slows down the process of copying because of the lower capacity. In order to be able to copy the next letter, the working memory capacity needs to be freed. A similar process occurs for words. If someone can write words, then it is assumed that their letter forming is already fluent, or has reached automaticity. Demands on capacity may be required for remembering some difficult words. If someone can process larger word chunks, then most of words can be copied automatically, without the need to use most of the resources in the working memory. The demands on capacity may then focus on the amount of words taken as one big chunk (e.g. Van Genuchten & Cheng, 2010).

An interesting study on copying was done by Weinzierl, Grabowski & Schmitt (2012) looking into the copying ability of second and fourth grade children in relation to working memory. The methodology uses four different types of symbol, which varied in terms of their semantic and phonological characteristics: geometrical symbols, unpronounceable consonant strings, numeral strings, and meaningful text. It was found that the copying of meaningful text was the quickest, followed by numerals, then consonant strings, and finally by geometrical symbols. This sequence was the same for both grades, and it is not surprising that fourth graders could copy more quickly than second graders, because of their level of competence and the nature of processing. Copying geometric symbols is similar to copying characters from an unfamiliar script. The copying of unfamiliar characters would of course cost more time than familiar alphabet letters.

#### **2.3.4 Sentence-Copying as a Predictive Task to Investigate the Cognitive Processes in Language Processing of ESLs**

In the writing research literature, writing processes are divided into two separate groups: high-level and low-level processes of writing. In recent years, there has been an increasing interest in the investigation of cognitive processes involving the high-level processes of writing, but very few involving the lower-level processes. A clear distinction was made between these two levels by Grabowski (2009), who stated that: *'while high-level or planning processes involve the generation, selection and ordering of ideas, low-level activities relate to lexical access and sentence generation (formulation), and to graphic transcription and execution'*. Therefore, it is possible to measure the cognitive processes by examining copying sentences (which is considered

a low-level process), which involve language components, such as word knowledge and grammar structure. The activity of copying will be captured through pause analysis, which measures the lengths of pauses between word chunks.

The ability to chunk a greater amount of words depends highly on the level of familiarity with the words, how frequently one sees the words, and the word length itself, which also acts as a proxy to measuring the components of cognitive processes in language processing in this study. Detailed explanation on word characteristics involving the notion of familiarity, frequency and length can be found in Section 4.2.2. Copying is a skill built through substantial practice and years of handwriting experience. Copying sentences fluently requires great familiarity with handwriting production, and word- and grammar-knowledge. The fluent execution of copying implies that a number of tasks become automatic, especially handwriting production in adults, who are the target participants for this research. Copying can be more fluent when the copier has greater knowledge of the vocabulary. When a person is competent in the language, the sentence can also be easily comprehended. More discussion on the notion of comprehension and how it relates to the cognitive processes is explained in Section 4.2.2.

As mentioned by Schilperoord (1996), there is a need to control the data captured by the pauses, so that they represent only the activities being investigated. Comparing the activity of copying with essay composition clearly demonstrates that the former can allow a higher level of control. For the purposes of this thesis, which intends to apply pause analysis in order to assess language competence, 'copying' assists in focusing on a fine level of writing (lower level processes of writing), where pauses can represent the level of understanding at word level processing.

## 2.4 Second Language Processing

Here, researches in relation to second language processing are reviewed, in particular those related to the cognitive processes.

In the world of second languages, terms such as multilingualism, bilingualism and monolingualism are used to categorise individuals into the number of languages they can speak. Multilingualism refers to the ability to use two or more languages. Some linguists and psychologists use bilingualism to refer to only two languages and multilingualism to more than two languages, but the issue of differences in terms is not our focus here. Monolingualism, then, refers to the ability to use only one language. The multilingual population includes nearly half of the world's population, but it is surprising that research has been focusing mostly on monolingual conditions and first language acquisition (Saville-Troike, 2006). Research involving second language is slowly increasing, as First Language (FL) and Second Language (SL) processes cannot be equated, nor can a multilingualist be assumed to have the same knowledge and skills as a monolingualist. With technology booming and new methods of research being found, it is becoming possible to explore, study and reveal the underlying processes, to investigate the differences of language processing between individuals.

Cook (1992), who highlighted the term *multilingual competence*, says:

*SL users differ from monolinguals in FL knowledge; advanced SL users differ from monolinguals in SL knowledge; SL users have a different metalinguistic awareness from monolinguals; **SL users have different cognitive processes.** These subtle differences consistently suggest that people with multicompetence are not simply equivalent to two monolinguals but are a unique combination.*

*(Cook, 1992: 557)*

Therefore, what affects the level of competence in a second language must be taken into consideration, such as the wide range of language learning settings, learner characteristics and circumstances. And of course, how one defines the factors may also be affected by the background of the researchers: linguists, psycholinguists, sociolinguists and social psychologists. This leads to the question: Why are some SL learners more (or less) successful than others?

### 2.4.1 A Glimpse of the Second Language Learner - How Learner Acquire SL Knowledge Impacts on the Cognitive Processes?

One aspect in finding out: '*why some SL learners are more (or less) successful than others*', may be related to how they acquire the SL knowledge (Rubin, 1981, 1987). Factors could be innate capacity, the application of SL prior knowledge to FL, the in-depth level of language processing involved, the degree of interaction in communication involving the SL, the restructuring of the SL knowledge system through the intensity of interactions with environmental factors which is motivated by FL knowledge and by input of SL, the mapping of relationships or associations between linguistic features, functions and forms such as grammar, and finally, the issues of automatisisation in processing capacity, especially when one has been exposed to frequency of input as well as good practise in processing input and output (e.g. Rubin, 1987; Saville-Troike, 2006; Wharton, 2000). These factors indirectly impact how the SL processes the sentence-copying, in terms of the size of chunks, which would then produce a different range of pause values.

Given an example of two SL learners, X reads many English books and Y is not keen on reading. The differences can be seen in how they chunk the words in the sentence and the pauses they produce. It is not surprising that X may have a large vocabulary knowledge from reading, hence have no problem with recognising words automatically. X is capable of producing large chunks of words. Y, on the other hand, may have a slight problem with comprehending the sentence, or understanding the meaning of words, hence would most probably produce smaller chunks. Because X is familiar with most of the words (the effect of reading many English books), X may produce larger chunks and shorter pauses overall. In contrast, Y may produce smaller chunks with longer pauses overall, because of the unfamiliarity of words. To conclude, the way SL is exposed to the language does impact on the cognitive processes and this is how one SL is different to another SL.

### 2.4.2 A Glimpse of Beginning SL Reading

Learning about the beginning of SL reading might assist in understanding how some learners may differ from other learners. For example, the priorities for SL activities differ with regards to academic competence versus interpersonal competence (e.g. Anderson, 1991). Academic competence defines learners who aim to use SL primarily to learn about other subjects, for research or occupation, and require reading as the most prioritised activity followed by listening-writing-speaking. On the other hand, interpersonal competence defines learners who aim to use



the SL primarily for communication use so prioritise listening, followed by speaking-reading-writing. Because academic competence do more reading compared to interpersonal competence (who prioritises listening), the amount of SL word recognition is predicted to be better with academic competence, which in this case of thesis work, would provide a different impact.

In addition, the fluency of reading, which is an essential aspect of academic competence, takes time to develop, in either FL or SL. Indeed, the level of FL reading is a strong indicator of how successful students will be in reading SL. This has been proven by Saville-Troike (e.g. 1984): even though the FL in their research was in a symbolic writing system (orthography), readers of Japanese or Hebrew were able to transfer their reading skills to English as a SL. Differences in language processing can still exist as there are variations in the degree of comprehension (see Section 4.2.2 for further explanation) because of the different syntax structure and vocabulary, even with FL learners.

Learners whose FL is written in a different orthographic system from their SL may need to familiarise themselves with the symbols of the SL as an early step (e.g. Stephanie, 2009). The different writing system in a different language not only involves the aspect of orthography, but also the syntax structure itself (Ehri & Wilce, 1980) (refer to Table 2.1). Therefore, it may not be surprising that these factors of orthography and mapping of FL knowledge to SL words may affect the successful level of learners. In the case of this thesis work, the same factors may also affect the pause lengths produced. Choosing SL learners with the same orthographic system may then assist the investigation in focusing on the components of cognitive processes in language processing.

<b>Alphabets</b>	English	She went to the market.	
	Lao	ເປັນຄົນໄປຕະຫຼາດ.	law pa:l thala:t she go market
	Greek	πήγε στην αγορά	pighe stin aghora he/she.went to.the market
<b>Syllabary</b>	Japanese	に行ってきました行に市場は彼女	kanojo ga maketto ni ikimashita She nom. market to went
<b>Logographic</b>	Chinese	她去市場	ta qu le shi.chang she go perfect market

Table 2.1 Writing systems of the world

Besides the issue of a different orthographic system, Table 2.1 provides evidence that different languages portray a different syntax structure, and therefore doing a direct translation may

somehow interrupt the process of chunking. This may then indirectly affect the pause lengths. Translating may assist with comprehending the sentence, but may not help in terms of how one chunks the group of words.

Other factors that may affect reading could be the components of language knowledge itself: vocabulary (lexicon), morphology (word structure), phonology (sound system), syntax (grammar) and discourse (ways to connect sentences and organise information) (Gass, 2008; Saville-Troike, 2006). Taking syntax again as an example, the ordering of subject and predicate in a sentence may differ radically between languages. Some ordering examples can be seen in Table 2.2. Nevertheless, these components are automatically available to children for their FL, and are usually acquired with no conscious effort. Conversely, even the most highly educated adult native speakers could not master all aspects of SL knowledge (Saville-Troike, 2006) because of a variety of reasons. These reasons may lead back to how the learners acquire the SL from the very beginning.

Subject Verb Object	English, Chinese, French, Russian
Subject Object Verb	Japanese, Turkish, Persian, Finnish
Verb Subject Object	Irish, Welsh, Samoan, Zapotec

Table 2.2 The order of syntax structure according to language

### 2.4.3 Why: Learner Differences?

The differences in learners of SL itself may have some impact on how they process the language, whether in reading, word recognition or comprehension. The factors that make one learner different from another may be as follows:

1. Gender issues in relation to cognitive style or learning strategies may have some effects on SL learner differences. According to Saville-Troike (2006), there is mixed research evidence in terms of language acquisition and processing. For example, Kimura (1992) found women to outperform men in some tests of verbal fluency, such as finding words that begin with a certain letter, and also that women's brains are less asymmetrically organised than men's in terms of speech. Halpern (2000) found, in relation to mental representations in vocabulary versus grammar, that females are better at memorising complex forms, while males are better at computing compositional rules. There are also some other studies on hormonal variables that affect female and male language processing and production, which can be seen in Mack (1992) and Kimura (1992). This gender issue may have some impact on the pause lengths produced. For example, females who can

memorise complex forms may be able to process jumbled words better, producing shorter pause lengths compared to males.

2. Learner differences may be affected by aptitude, in which there are assumptions that language learning may be based on talent. Carroll (1965) proposed four components underlying this talent: phonemic coding ability, inductive language learning ability, grammatical sensitivity and associative memory capacity. Although these components may not be completely deterministic as a predictor for differential success of SL learning, they form part of several factors that may influence SL proficiency (Skehan, 1998). These four components may affect how one processes their language, and this can be seen by the variety of pause values seen for one individual (doing several sentences) and even amongst a group of SL learners.

3. Cognitive style, which refers to the preferable ways of processing, i.e. of perceiving, conceptualising, organising, and recalling information (Rubin, 1981; Saville-Troike, 2006). Although this has not been as well established as factors such as the issues mentioned above, cognitive style itself is closely related to and interacts with personality factors and learning strategies (Saville-Troike, 2006, p. 87). In this thesis, the cognitive style would of course have some effect on the chunking of letters and words in the sentence-copying task. It is undeniable that one's cognitive style influences the way one processes the language, which is what makes all the pauses produced unique to an individual person. It is rare that the pauses produced are similar to another individual.

4. The fluency of reading, as reviewed by Grabe (1991) may also affect the processing of language, namely:

- Automatic recognition ability
- Vocabulary and structural knowledge
- Formal discourse structure knowledge
- Content/world background knowledge
- Synthesis and evaluation processes/strategies
- Metacognitive knowledge and comprehension monitoring.

Alderson (2000) claimed that readers can process the written form much faster than sound, and that a fluent reader could read three times faster than the speed of normal everyday speech. This

would mean that, if someone were to translate every word they read into sound, then this would increase the time spent reading threefold. The purpose for reading also influences the time spent in the reading process. Time spent searching for information will be longer than time spent skim reading. Grabe and Stoller (2002) categorised the different purposes of reading into: searching for simple information, quick skimming, learning, integrating information, writing, critiquing, and basic comprehension. It could also be the case, however, that the copier may not be highly competent in the language but simply naturally writes quickly, and this could affect the data. Avoiding the inclusion of such individuals in the experiment does not solve the problem, but there is certainly a need to tackle this particular case during the experiment investigation.

The ability to read fluently also depends on two terms that are normally included in relevant discussions: skills and strategies. Skills are normally related to the linguistic processing abilities that are relatively automatic (e.g. word recognition, letter recognition), while strategies are sets of abilities that are consciously controlled by the reader. There are debates around these two terms. Some theorists have argued that strategies, when practiced serially, could become as automatic as skills (Grabe & Stoller, 2002). Skills, however, could also become strategies when used intentionally (Paris et al., 1991; Alexander & Jetton, 2000). As a rule of thumb, theorists have defined skills as automatic processes and strategies as abilities based on conscious reflections. Strategies could also be called '*skills under consideration*' (Paris et al., 1991). In relation to sentence-copying, there is a high possibility that copiers will adopt a copying strategy. The experiments will explore further whether a control is needed or not.

#### 2.4.4 How?: The Learning Processes

What makes a learner more (or less) successful than another, other than the learner differences, may be related to the learning processes involved. Psychologists provide two major frameworks related to learning processes: **Information Processing (IP)** and **Connectionism**. IP, as developed by Anderson (e.g. 1976, 1983), claims that learning a language is like learning any other domain of knowledge, where one does not engage in any other mental activity, but focuses only on the learning. Connectionism as a framework is similar to IP but is claimed to require parallel processing and focuses on the increasing strength of associations between stimuli and responses, rather than the abstraction of rules or principles.

IP has three stages, consisting of input (perception), central processing, and output (production). Input means whatever sample an SL is exposed to (e.g. reading materials); this input then needs to be noticed by the SL in order to be processed. . In the case of this research where a copying task is applied, the input is controlled, which forces learners to notice the items given. The degree of noticing or awareness of input (Schmidt, 1990) provides us some chance of differentiating the participants, even with a simple activity such as copying, which only considers low-level cognitive processing. These features which act as contributors to the level of noticing are:

- Frequency of encounter with items
- Perceptual saliency of items
- Instructional strategies that can structure learner attention
- Individuals' processing ability (a component of aptitude)
- Readiness to notice particular items (related to hierarchies of complexity)
- Task demands or the nature of activity the learner is engaged in.

Central processing, where learning occurs, is the heart of this IP model. It is when (i) controlled processing improves to an automatic processing and (ii) restructuring takes place that makes mental representations more coordinated, integrated, and efficient. For example, one research found that the amount of time taken for a multilingual to recall words and grammatical structures shows that the SL of even fluent speakers of both languages is generally less automatised than the FL, and less proficient SL is less automatised than more fluent SL. Certainly, this central processing requires practice in order to become more automatised (Saville-Troike, 2006).

The third stage, the language output in which the SL produces, may be in the form of spoken, signed or written language; has been a profound finding of Merrill Swain (e.g. Swain and Lapkin, 1995). Swain (1985) who focuses on the importance of successful output for SL learning emphasises that doing lots of practice would assist in developing the automaticity of the fluency of the language. Automaticity seems to be one of the success ratings for language learners.

As a whole, the conclusion that can be made from the two models, IP and Connectionist, is that both models entail 'frequency' and the level of 'attention' given to the language as being what influence language learning. These factors, which affect the automaticity of processing the language, are claimed by the SLA researchers to be better than doing repetition drilling. For that reason, it is not surprising that the processing of language differs from one to another because of

the different amount of exposure to the SL (i.e. frequency) and how engaged the learner is: what level of attention they give to the details of that particular knowledge.

## 2.5 Factors Affecting the Cognitive Demands of Copying: The Processing Costs

This section explores the factors that might contribute to the demands of ‘processing costs’, which affect pause lengths. In this thesis, ‘processing costs’ relate mostly to handwriting production and word- and meaning-related processes (which will be covered in this section), as they are included in the lower-level processes of writing. The word- and meaning-related processes relate to the different level of exposure the SL has to the language and one must bear in mind that there are many factors that differentiate the success level of learners. All three main factors are very closely related to the phenomenon of ‘automaticity’. *‘When a component skill is automated, it becomes faster, effortless and non-interfering’* (Fayol, 1998) and hence does not overload the limited capacity of memory. It thus allows participants to devote more resources to activities that are not automatic.

In terms of measuring language, the so-called ‘processing costs’ will depend on the individual’s level of language competence. To illustrate, the more proficient a writer is graphomotorically (i.e. in his/her graphic transcription processes), the more automatised the handwriting production becomes, and the more space is available in the working memory for other processes, such as retrieving spellings, words and meanings of the processed language from memory (Bourdin & Fayol, 1994, 1998; McCutchen, 1996, 2000; Bonin, 1997; Bonin et al., 1997; Graham et al., 1997; Graham & Harris, 2000). To demonstrate further, a competent user with a better understanding of the language might have reached a high automaticity level when it comes to the retrieval of familiar words and meanings from memory, and thus require less processing of resources, which results in shorter pause lengths. Someone less competent with the language, on the other hand, might be slower in understanding the stimuli, and take longer to retrieve familiar words (including spellings) and meanings from memory, hence demanding more processing resources, so producing longer pause lengths.

To conclude, this section explores the relevant literature and proposes some factors that might contribute to the demands of processing costs during the activity of copying. This would also include the role of working memory and the effects of English as a second language.

### 2.5.1 Working Memory Related Processes

Working memory, understood as a model of short term storage and processing, has also been used to study and understand cognitive processes, demands and capacity. Language processing like writing demands a lot of cognitive processes, for example recalling words from long term memory and structuring sentences. This cognitive process demand varies depending on the expertise level. Kellogg (1998) mentioned '*writers with a greater knowledge of a subject would require less cognitive resources compared to less knowledgeable writers*'. This finding can be mapped on to the prediction of the fluency in copying English sentences in this research, where a more competent English language user is predicted to demand less cognitive effort compared to a less competent user, hence the more competent user would be likely to have shorter pause lengths than the less competent user.

#### 2.5.1.1 The Role of Working Memory

The working memory modelled by Baddeley and Hitch (1974) has three main components: the central executive, the phonological loop, and the visuo-spatial sketchpad. The central executive is the main controller, served by two slave subsystems: the phonological loop and the visuo-spatial sketchpad. For example, in the activities of copying words, all three are used but the degree of use depends highly on the competence level of that particular individual. It is clear that the visuo-spatial sketchpad which processes the visual and spatial information is highly involved in the copying processes, but the phonological loop may or may not be used. The copying applied in this research does not involve speech processes, although there may be times where subvocal rehearsal may be used during the copying. At times, signals from short term memory may be sent to the long term memory to search for word meanings (i.e. semantic processing).

Cognitive effort or demand is likely to be reflected in the pause value, but, we need to know how this relates to the processes of copying the sentences and the processes in the working memory. How does a writer process the copying of a sentence? What is a pause? To answer these questions, it is important to examine the detailed processes in the working memory, such as the recognition of letters or words, and the number of letters or words held by the short term memory (STM). The process of copying may also involve the ability to relate the words to each other in order to build up the whole sentence meaning and the ability to reproduce the words by writing. On top of that, the copier may or may not refer to the semantic system in order to search for word meanings; this depends on one's competence level of that person. Each of these

processes takes some time, which is what we refer to as the pauses. These differ in length between individuals, depending on their level of knowledge in that particular language. A long pause at any one point in the process of copying may represent a greater demand of working memory usage. For example, someone may be struggling to recognise a low frequency word, such as 'autonomous', and therefore would be likely to chunk the word into smaller groups of letters to ease the copying process. Hence, the involvement of working memory is substantial in the process of copying because of the processes that allow copying to take place. The existence of a relationship between working memory and sentence copying provides us with a means to assess language competence.

#### *2.5.1.2 Mental Storage Capacity: The Study of Chunks*

When copying, the writer needs to read and store as much information as possible in their short term memory. What is stored may consist of a number of words, syllables or letters. The ability to store information depends on one's memory storage capacity, and the total number of words that someone can store depends on how they group the words, syllables or letters: what Cowan (2001) and Miller (1956) refer to as the 'chunking process'. According to Cowan (2001), *'a chunk must be defined with respect to associations between concepts in long term memory'*. He then defined the term 'chunk' as *'a collection of concepts that have strong associations to one another and much weaker associations to other chunks concurrently in use'*. It is assumed further that the number of chunks can be estimated only when inter-chunk associations are of no use in retrieval in the assigned task, i.e. when there is no relation of meaning between chunks. Cowan (2001), in his studies, has come to the conclusion that the capacity limit is an average of four, fewer than Miller (1956) claimed, which was  $7 \pm 2$  chunks.

The decision to chunk may be affected by the knowledge of phonology, graphology, orthography and language background, and whether it is the individual's first language or a second language. Short-term memory cannot hold too many pieces of information at one time, and information will slowly decay as other activity takes place. At this point, familiar words can be copied easily, but copying unfamiliar words may involve a different number of chunks or segmentation within a word. Since this research involves second language learners, it is highly probable that the chunking process would be affected by their first language. Therefore, the ability to chunk the sentence into groups of words, individual words, groups of letters or individual letters not only is determined by their exposure to English, but also involves their first language knowledge.



A famous example inspired by Miller (1956), gives the word **fbicbsibmirs**, which consists of well-known US associations' acronyms of letter triads in sequence: FBI, CBS, IBM and IRS. Someone who notices this would be able to use this information to assist recall. The same concept applies at word and sentence level. For example, a proverb may be easier to copy than a garden path sentence for many second language learners, because they are more likely to have been exposed to that proverb and therefore be familiar with it.

Given that the researcher herself is a second language learner, and that she is not familiar with some of the US associations named, she would have grouped the letters by associating them with a familiar group of letters or words that she has come across in her first language (Malay). As such, she would have chunked: FBI, CB, SIB and MIRS. FBI is a well-known US Government agency, seen on television. CB is chunked because these letters are next to each other alphabetically, even though B actually comes before C. SIB is part of the word 'NASIB', which means LUCK in Malay, and the syllables are usually divided as such, to assist spelling: NA+SIB. And finally, MIRS is because it is easy to pronounce MIRS like a word, although it might be meaningless.

In addition, there may be a possibility that chunking limits depend on the capacity limit of the *focus of attention* (Cowan, 1995; 2001). Cowan (2001), in his study of deliberate memory retrieval within a psychological task (e.g. recall or recognition) found that the retrieved chunk resides in the *focus of attention* at the time immediately preceding the response. This is also evidenced indirect memory tasks involving explicit memory such as recognition and recall: these require attention to the stimuli at both encoding and retrieval of information (Cowan, 1995). Similar attention, however, is not required in retrieval of information in indirect memory tasks, such as word fragment completion (for a review, see Cowan, 1995). This means that any information that is *deliberately recalled*, whether from a recent stimulus or from long term memory, may also affect the capacity limit of the focus of attention at that particular time.

### 2.5.1.3 Cognitive Capacity Constraints

One way of understanding how cognitive capacity works is to imagine it as a fluid resource shared amongst the mental processes. In the case of copying activities, motor planning of handwriting by a competent writer would run fluently, making very few demands on the cognitive capacity. However, when the writer encounters an unfamiliar low frequency word, copying the word would require higher demand processes because the writer is not familiar with the spelling or meaning.

The performance of the copying can proceed without any problem as long as the total demand does not exceed the available capacity for that particular individual. When total demand exceeds the limit, the copying production will be slower, hence the long pause.

Cowan (1995) suggest two processing limits. (1) Information in a temporarily heightened state of activation, but not in the current focus of attention, was said to be time-limited. (2) The transfer of this activated information into the focus of attention was said to be rate-limited. Cowan (1988; 1995) stresses further that only the focus of attention is assumed to be capacity limited. The focus of attention is the part that holds the current information, and the question is of how much information it can hold at any one point in time. This relates to the chunks size. The capacity sometimes depends on the kinds of tasks assigned to it, although there is no doubt that only a small amount of information can be kept consciously at any one time in this focus of attention (James, 1890).

Although the application of sentence copying may sound simple, the processing system can be overloaded by the writer being given difficult sentences to copy, such as unfamiliar proverbs, by jumbling letters within a word, and by using low frequency words. These difficult sentences may be a way to assess participants' nature of language processing, because the difficulty may push their cognitive demands to the limit, and will then be able to provide interesting information with regards to their levels of competence.

On the other hand, a competent person who has had good exposure to the language is able to develop domain-specific memory management strategies in order to make good use of available capacity. This will then allow them to retain information whilst copying and reading (Torrance and Galbraith, 2006). Moreover, one who is competent will be able to do certain things automatically (Torrance and Galbraith, 2006), for example in the spelling of high frequency words. There is no need to invoke long term memory in order to retrieve the spellings of certain words. For example, one could avoid stopping to clarify whether there are one or two 'c's in the word 'accessories'.

### 2.5.2 Handwriting Related Processes

It is a well-known fact (van Galen, 1991; Fayol, 1998; Longcamp et al., 2005) that handwriting is a complex skill that requires years of practice, but becomes automatised once mastered. The automatised handwriting processes contribute to the freeing of space in the working memory, which can then be used for language related processes. Since the aim of this thesis is to

measure the cognitive processes of language processing, it is important to make sure that the processing costs are free of any handwriting related issues. This section will briefly investigate the literature in relation to handwriting and how this is associated with cognitive processes.

#### **2.5.2.1 Studies in Handwriting**

Early research in the area of handwriting was mostly educational, concentrating on the developmental aspects of handwriting acquisition and the ergonomical features of handwriting techniques (Søvik, 1975; Barbe et al., 1984). From the 1960s to the 1980s, research in handwriting was mainly concerned with product-based topics, ranging from biomechanical conditions to the legibility of handwriting (Herrick & Okada, 1963; Askov et al., 1970; Peck et al., 1980), but none looked into language assessment. In the 1980s, a shift was observed from a product-oriented to a process-oriented approach in studying handwriting (Thomassen et al., 1984; Kao et al., 1986; Plamondon et al., 1989). It was not until well into the 1980s, however, that this approach really took off (Denier van der Gon & Thuring, 1965; Vredenburg and Koster, 1971; Wing, 1978). The same shift was also observed in the research of writing (e.g. Hayes & Flowers, 1980) and language (c.f. Alderson & Banerjee, 2001). One of the reasons behind this change of focus might be traced through the advent of technology and the growing application of computers, which have rendered many types of research possible. It is indeed true that it was the advent of electronic digitiser tablets that enabled researchers to investigate the real-time dynamics of handwriting production (e.g. Teulings & Thomassen, 1979; Teulings & Maarse, 1984). Further, a number of studies in the areas of latencies and movement times in handwriting tasks (e.g. van Galen & Teulings, 1983), neuropsychological observations on disturbances of handwriting related to localised brain lesions (e.g. Ellis 1982, 1988) and mathematical models of trajectory formation (e.g. Hollerbach, 1981) have contributed to the understanding of cognitive and motor processes in handwriting production.

#### **2.5.2.2 Handwriting & Cognitive Processes**

Van Galen (1991) introduced a model of handwriting production that explains the parallel and concurrent processes in handwriting and involves both biophysical and psychological aspects. He describes handwriting as being typically a *'compound, cognitive and motor skill'*. In his paper, *'Handwriting: Issues for a psychomotor theory'* (van Galen, 1991), he explains how writing is associated with semantic and lexical knowledge. More specifically, it relates to the language-specific sets of phonemes and graphemes of the letters of the alphabet, the different allographic

forms of upper and lower case letters, the mastery of alphabet knowledge for spelling purposes and the requirement for the delicate manipulation of the pencil in letter formation. Even though these aspects have already been researched, what is important for this thesis is the ability of handwriting to achieve the aforementioned processes and have them take place automatically.

There have also been consistent findings in relation to the spatial and temporal parameters of writing scripts, where handwriting is presented not only as contributing to the biomechanical conditions of a given course (e.g. movement time, writing size, writing fluency) but also as embodying the cognitive and motor demands of the task (e.g. van Galen et al., 1986, 1989; van Galen, 1990). The increase of movement time and trajectory length has been translated as the *'sharing of processing resources between real-time stroke production processes and concurrent preparatory processes concerning the forthcoming task segments'* (van Galen, 1991). This could suggest the occurrence of parallel processing, where multiple processes occur simultaneously, in the course of copying.

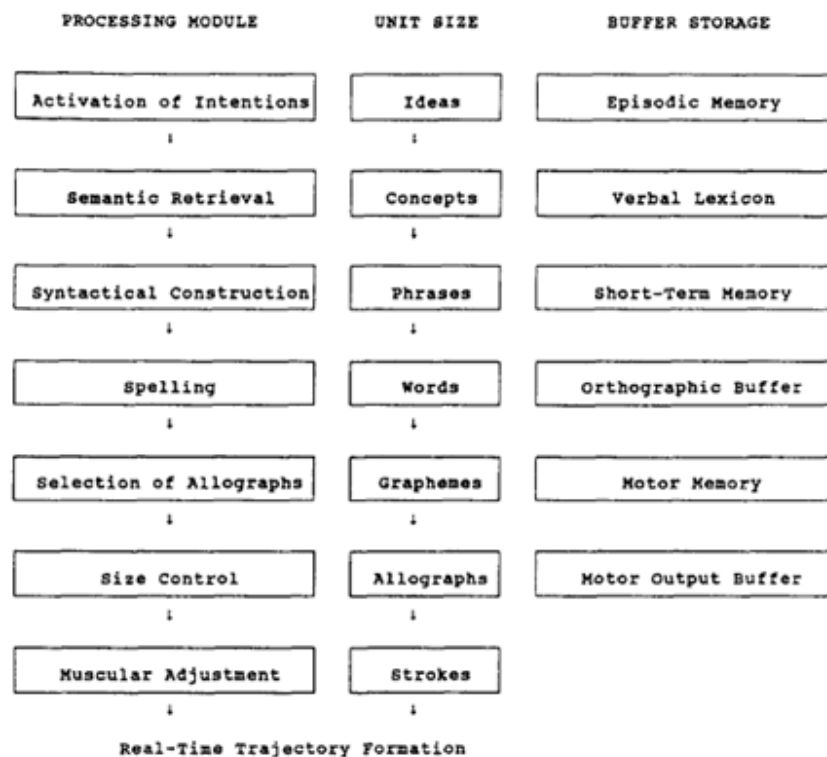


Figure 2.1 Handwriting Production Model by van Galen (1991)

NOTE: In the left-hand column, the hierarchy of processing modules is indicated. The central column describes the identity of the processing units addressed in the corresponding module. The

right-hand column refers to the storage nodes that mediate in the communication between successive levels of the model.

Figure 2.1 presents a model of Handwriting Production (van Galen, 1991), organised in a hierarchical manner and grouped according to processing module, processing unit and buffer storage, all of which run parallel to each other. This model has been used by Lambert et al. (2008) and Kandel et al. (2009) in order to understand the effects of syllables on spellings and handwriting production. The first two processing modules are involved in higher-level processing, usually employed in essay composition or text production. Particularly relevant to the purpose of this research are the lower-level stages: 'phrase' units and below, i.e. the lower order modules like words, allograph selections, size controls and muscular adjustments. This model assists the present research in distinguishing between the different processes that exist in the handwriting production model.

Working on the assumption that adults have achieved automaticity of handwriting (graphemes, allographs and strokes), the model focuses on phrase and word units, which relate closely to the assessment of language. Manipulation of the word forms (i.e. spelling irregularities) and of sentence meanings (i.e. syntax structure) would affect the buffer storage processes and it would hence take longer for signals to be sent to the motor modules. These increases pause lengths. The familiarity of spellings, word shapes and forms (i.e. orthographies) and word meanings contribute to the fluency of copying. Unfamiliarity would influence pause lengths negatively.

To conclude, the automaticity of handwriting is important in that it frees up space and resources in working memory. In order to make sure that processing costs are free of any handwriting issues, this thesis conducts experiments with a population group that comprising only adults, who spoke English as a second language, but whose first language also used the Roman alphabet. This enables this research to focus solely on word and meaning related processes.

### **2.5.2.3 Pressure on Handwriting**

There are also factors related to the stimulus used that may affect pause lengths because of the pressure given to the handwriting production. The different kinds of stimuli (*task difficulty*) used may affect the production of natural handwriting, known as *handwriting pressure* (Stelmach & Diggle, 1982; Kao, Shek & Lee, 1983; Kao, Mak & Lem, 1986) i.e. pressure on the handwriting process because of the unfamiliarity of the stimuli presented increases the handwriting production

time. Handwriting pressure contributes to the increase in processing demand in working memory (Kao, Mak & Lem, 1986). This might occur during the introduction of writing in the grid boxes, where one individual letter must be written in each box.

The researcher predicts that the use of boxes might affect natural hand movement when performing free handwriting, as writing in boxes is not as common practice as writing on a straight line. Barone (work in progress), however, finds that there is an increase of roughly 200ms between writing stimuli in boxes and writing on a straight line, and that this is observed throughout the copying activities. Hence, it can be generalised that these effects do not confound the approach, as the increase is fairly distributed across all writers and sentences used. Furthermore, because writing in boxes is unusual for all copiers, *'initial adjustments'* might take place in order to stabilise the writing activities, as suggested by Stelmach and Diggles (1982). Hence, the experiment introduces baselines before the real sentence-copying takes place. It is therefore assumed that when the copier starts copying the sentences, handwriting production will have become more stabilised and automatic. Further manipulations on the sentences that increase the difficulty of copying do not, however, affect handwriting pressure, because handwriting production is automatic for adults. This should mean that the pauses represent only language-related processes.

### 2.5.3 Word Related Processes

This section focuses on the word related processes involved in the activity of copying sentences. These include the familiarity level of word recognition, word knowledge, and spelling, as well as the understanding of the meaning of the words. In fact, English, known as a 'deep language', is problematic to work with, as it contains *'deep orthographies that can be difficult for some to internalise'* (Grabe & Stoller, 2002). The familiarity level of each participant is unique and, as such, cannot be determined in a straightforward manner. When the words are very familiar, the recall process from long-term memory would be quicker than with less familiar words. For instance, when participants read a word but are unsure of the spelling, the word will be stored in the short-term memory, which will send signals to the lexical database to search for correct spellings. This will result in a longer process, than that put in place in response to familiar words. Therefore, this part of the chapter will examine the factors related to word processing, which could affect the pause lengths of each participant involved.

### *2.5.3.1 Studies in the Visual Word Recognition Model*

A number of researchers have worked towards the development of models to study word recognition. The impetus behind their research was mostly to address reading difficulties, for example in dyslexia. Nevertheless, these models are instrumental for the present research in terms of understanding the processes involved in word recognition. Marshall and Newcombe (1973) pioneered a cognitive model of reading, also known as ‘visual word recognition’ that relates to the processing of words, be they proper words or non-words (i.e. have no assigned meaning), (e.g. van Order, 1987; Ellis & Young, 1988; Taft, 1991, 1994; Coltheart et al., 1993, 2001; van Order & Goldinger, 1994, 1996; Plaut et al., 1996).

Further, the computational model of reading that is the most referred to in the field of reading and visual word recognition, and is considered the most useful, is the Dual Route Cascaded (DRC) model (Coltheart et al., 1993). This model explains the processes of reading words aloud. Coltheart based his work upon that of Patterson and Shewell (1987) and their model of language processing. A number of researchers interested in investigating the linguistic features of words have used the DRC model in order to understand the processes at syllabic level, e.g. the work of Verhoeven et al. (2006), Kandel et al. (2006a, 2006b, 2009) and Lambert et al. (2008) (cf. Section Syllables Processing). Both these models make use of the ‘phonological output lexicon’, which does not relate directly to the aims of this thesis. The present research, however, aims at looking at the printed form of language as both an input and an output.

Alderson (2000) reported that theorists have not yet agreed on the skills involved in reading processes. Hence, researchers tend to make assumptions of what actually takes place during these processes. For example, Alderson (2000) asked whether readers relate the printed form of language with the spoken form of the text. As demonstrated by Patterson and Shewell (1986) in their proposed language processing model, reading aloud might involve translating print to speech. The findings by Kandel et al. (2006), however, reported that, given a list of words containing both orthographically and phonologically driven examples, writers copied all words based on orthographical conventions, ignoring the phonological ones. Another argument by Smith (1971), which relates to that of Kandel et al. (2006), was that readers proceeded directly to the meaning and were not influenced by sound. Similar conditions are expected from the sentence-copying processes.

### 2.5.3.2 *Recognising Word Shapes*

Orthography is a system of marks used for representing spoken language in writing. For some people, recognition of word shapes is built upon their orthography knowledge. When a reader is very familiar with the language, a missing alphabet letter would make the shape of a word look odd; this would particularly affect *sight word reading*. The same would occur when an extra alphabet letter is inserted or when letters are in an incorrect position. This level of familiarity signifies that the recognition of words has reached the level of an automatic process. Any transposition of letters or missing letters in a word will slow down the reading process, hence increasing pause lengths. Factors such as letter knowledge and phonemic awareness also contribute towards recognising the word shape (e.g. Ehri, 1994; 1998; 2000; 2005).

Working with ESL adult participants, however, letter knowledge and phonemic awareness is assumed to be an automatic process, even though there is always the possibility that less competent ESL participants might resort to the strategy of copying new English words with their first language (FL) spelling methods. Even so, the amount of practice and exposure they might have had with English printed materials might also affect and determine their competence level.

### 2.5.3.3 *The Ability of Sight Word Reading*

Sight word reading is described by Ehri (2005) as the process whereby a '*reader reads familiar words by accessing them in memory*'. With practice and experience, a reader can read words automatically by sight. An increased amount of word knowledge or vocabulary helps a reader to process more advanced text. Less practice and exposure to certain words makes the process of accessing them from memory slower and hence increases the processing demands, which affect pause lengths, identifying them as less competent users of the language. The process of learning sight words involves the ability to spell the words, '*forming connections between phonemes and graphemes to bond the spelling of the words to their pronunciations and meanings in memory*' (Ehri, 2005). In the case of letter transposition, however, (e.g. Rawlinson, 1976; Christianson et al., 2005), the process of sight reading would be difficult and thus affect pause lengths.

In brief, the possibilities of sight word reading occurring during sentence-copying depend on various conditions. There is no doubt that a competent person is able to sight read fluently. Manipulation of spellings in the sentences, however, might increase pause lengths. This is because the manipulation would change the word shape, even if the word is still recognisable by the letters that it holds (e.g. Christianson et al., 2005; Sears et al., 2006; Velan & Frost, 2007). Somebody who



is less competent might only be able to sight read high frequency words that are very familiar. In addition, level competency in sight word reading also affects chunking size; i.e. the number of words in a chunk.

#### *2.5.3.4 The Effects of Syllables Processing*

To become a 'conventional speller' (as described by Moats, 2005), especially in English, one must learn the orthographic system in English, namely that graphemes, the units that represent phonemes, often contain more than 1 letter, e.g. /th/, /sh/, /ch/ and that many vowel spellings are actually vowel teams (igh, ei, aw, oy) (Moats, 2005). In addition, there are several other rules one must be aware of, like recognising endings (morphemes) added to words (e.g. -ing, -ly), recognising meaningful parts, such as compounds, and being acquainted with the grapho-syllabic conventions (e.g. the use of doubled letters after short vowels). Whilst mastering the conventional spelling, students of ESL learn to store the letter patterns in memory in the form of "chunks"; i.e. syllable spellings, common endings and word parts, and high frequency words. These syllable chunking forms might be a strategy used to copy words.

A group of psychologists recently explored the role and effects of the number of syllables in a word on spelling (Kandel et al., 2006a, 2006b & 2009; Lambert et al., 2008). Their study showed that the syllables in a word affect the process of copying different words in Dutch and in French. Factors were manipulated, such as the length and number of syllables in a word. Kandel et al. (2006a, b) also examined the different processes involved in phonology and orthography and showed that copying is not influenced by the phonology of a word but is instead orthographically processed. The number of syllables in a word, on the other hand, does affect pause lengths during copying. For example, the longer the word (e.g. in-de-pen-dent), and hence the greater the number of syllables it might have (in this case four), the more time needed for processing (because there are three positions of pauses within the word) and hence the production of longer pauses. Less competent participants are more prone to such phenomena. By contrast, a very competent participant might be able to copy the word 'independent' as one big chunk with a long pause at the beginning of the word.

Verhoeven et al. (2006) introduced the Extended Dual Route Cascaded Model (Figure 2.2) based on the standard DRC model (Coltheart et al., 2001). Verhoeven et al. (2006) proposed a graphotactic bypass that might be used by some writers, as it is governed by visual word forms.

The model also introduces a ‘segmentation’ process, which allows the chunking of words based on graphotactic rules, which could occur in the processing of difficult words. A similar model was also introduced for the processing of morphologically complex words (Schruefer and Baayen, 1997). In relation to this thesis, however, this research does not consider processes that involve phonology, because it is assumed that in the process of copying, copiers use the orthographic strategy. This is supported further by the findings of Kandel et al. (2006a, 2006b, 2009) who have shown that, even though words are pronounced differently phonologically, the process of writing is still determined by their orthographic features. Hence, this research will ignore the ‘phoneme system’ and the output of ‘speech’. Mapping the findings of Kandel et al. as mentioned above to the working memory model, the copying approach that this thesis focuses on may ignore the use of the phonological loop processes that involve sound, and instead focus on the visuo-spatial system which processes these words according to their visual perception, i.e. orthography. A similar structure to this model is adopted in the theoretical model for understanding the copying processes developed and presented in Chapter 4 of this thesis.

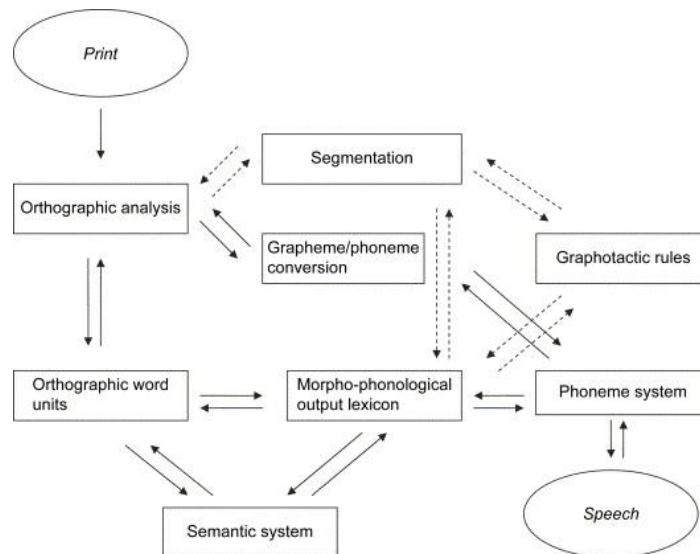


Figure 2.2 Basic Architecture of the Extended Dual Route Cascaded Model (dotted lines show the extensions as compared with the standard DRC model) (Verhoeven et al., 2006)

### 2.5.3.5 The Effects of Word Lengths

Word length always varies in a sentence. Processing words that are short is easy for all levels of competence. Long words, however, like ‘*intermediary*’ would be difficult for a less competent English user. A study by Baddeley et al. (1975) agrees that as the length of the words gets bigger and as the number of syllables increases, then recall performance becomes poorer. A person who

is very competent in English and has good vocabulary knowledge would most likely be able to copy the whole word fluently, just by sight reading. In terms of pause studies, fluency is defined as having a fairly uniform distribution of pause value between each letter (L1 letter level pause) in the word. For example:

'i(78)n(82)t(80)e(75)r(70)m(69)e(80)d(75)i(70)a(70)r(65)y'.

The numbers in the brackets represent pause values and this is an example of pause values being distributed almost equally. The same phenomenon, however, might not apply to a less competent English user. A group of letters might be chunked together, maybe at syllabic level, or maybe in a manner influenced by the individual's background (e.g. experience and culture). Chunking at syllabic level can be observed when a longer pause (longer than the pause values between letters) occurs after a group of letters (see in **bold**). An example of this would be:

'i(82)n(**180**)t(70)e(65)r(**225**)m(60)e(**175**)d(68)i(69)a(**110**)r(60)y'.

A longer pause value can be observed between 'in', 'ter', 'me', 'dia' and 'ry'. However, between 'dia' and 'ry' there is a much shorter pause compared to the other three, and this may be due to the fact that this particular person is familiar with the word 'diary'.

Individuals each have unique strategies of copying, which are affected by many factors, as explained above and in previous subsections. These factors obviously affect pause length, which represents the cognitive processes involved. In certain cases, parallel processing might also occur, in which a number of processes occur simultaneously. For example, eye movement when reading is happening as the hand is still writing. Eye movement is quick when words are familiar, especially when sight reading is fluent. This might be one of the factors that differentiate between competent and the less competent users.

#### 2.5.4 Meaning Related Processes

The ability to decode letters and pronunciations of letter strings does not guarantee the successful achieving of language comprehension. As argued by Scarborough (2001), *'text will not be well comprehended if one (1) does not know the words in their spoken form, (2) cannot parse the syntactic and semantic relationships among the words, or (3) lacks critical background knowledge or inferential skills to interpret the text appropriately and read between the lines'*. Moreover, characteristics of a text, such as the cultural background of the topic and the level of

difficulty of the vocabulary, were also found to have an impact on reading comprehension (Johnson, 1982).

From the perspective of pause studies, the ability to organise information in large chunks, namely with many words and shorter pauses at the beginning of the first word in that particular chunk, would define those who are competent, demonstrating an automatised language comprehension process. By contrast, smaller chunking processes at word level with longer pauses (used for trying to understand the meaning of the words) or even at the syllable level with many short pauses (because of the lack of vocabulary knowledge), would likely define a less competent language user. Various theories and models have been introduced in the attempt to understand what is involved in the process of reading comprehension (e.g. Goodman, 1967; Gough, 1972; LaBerge and Samuels, 1974; Kintsch & van Dijk, 1978; Just & Carpenter, 1980; Rayner & Pollatsek, 1989; Perfetti, 1999; Cohen & Upton, 2006; Birch 2007), some of which were introduced in the previous section. From the literature, three factors in relation to meaning related processes were identified: background knowledge, vocabulary, and language structures, all of which are assumed to affect the cognitive demands of copying.

#### ***2.5.4.1 Background Knowledge of the Subject Under Discussion***

Background knowledge of the subject discussed is important in terms of the familiarity of the words used in a sentence. As a simple example, if the subject specific knowledge of the copier is law, then it would be difficult for them to understand and process sentences that contain terminology from mathematics or chemistry. The level of familiarity with the terminology used is low, thereby affecting the process of comprehension, resulting in long pauses at L2 word level, which represents semantic processing. Similar studies have been conducted, for example Taylor (1979), which concluded that there are no differences between good and poor readers when it comes to familiar texts, but poor readers are significantly less able to recall unfamiliar text. Good and poor readers in this context simply reflect the amount and type of reading that they have been exposed to (Taylor, 1979). Equally, Recht and Leslie (1988) did a study on 'recall of text' and concluded that when the reader has greater knowledge of the subject, comprehension is better and fewer errors are produced when text is recalled. The same observations are expected of the copier who has greater subject specific knowledge matched to the sentence stimuli.

#### 2.5.4.2 *Vocabulary Knowledge*

It is now known that word recognition is important for fluency in reading (as explained in the previous subsections). Knowing how to sound a word, however, is not enough. One must know its meaning in order to be able to read it with better chances of comprehension. Fluency in reading is arguably influenced and affected by word *forms* and their *meanings*, also known as vocabulary or lexical competence. Vocabulary development is important for reading comprehension (Johnson, 1982). Strong correlations between the knowledge of word meanings and the ability to comprehend a text when words were made easier to understand were found in previous studies (e.g. Wittrock et al., 1974, 1975; Anderson & Freebody, 1979; also refer to Section 2.2.4). There is also a growing body of literature investigating vocabulary measures, such as the work of Meara (1996), Laufer (2006) and Nation (2007), but, again, these measures use the offline testing approach.

In relation to this study, good vocabulary knowledge is heavily related to the previous factors discussed, such as being able to sight read quickly, chunk larger groups of words, and process meaning quickly. All of these affect pause lengths; they mean that shorter pauses are produced, as the person does not need to refer to the semantic system and can therefore process the word automatically and at the same time as the hand is producing the written form. This also references parallel processing as mentioned above. The pause analysis approach might also contribute to the study of vocabulary measure, as it offers an online test for investigating the cognitive processes underlying vocabulary competence.

#### 2.5.4.3 *Language Structures (Including Grammar)*

Sentences that contain a difficult structure might also affect comprehension. A sentence can be passive or active, and both of these states contain a level of difficulty in understanding (e.g. Miller, 1962; Mehler, 1963; Miller & McKean, 1964). Sentence production can also be affected by the writer's cultural background and their first language. Even though most learn the basic structures of sentence production in school, there are various factors, such as writing styles and vocabulary choices that affect the comprehension of a text. Sentences like proverbs are sayings whose familiarity can depend on cultural background and exposure to the proverbs itself. Garden path sentences, on the other hand, are sentences which are deliberately misleading, and can be easily misinterpreted unless they are read carefully (e.g. Lewis, 1993). Competence in comprehending these types of sentences depends on the familiarity level of the reader, both with cultural and

with grammatical elements of the language. If presented to a less competent reader, they might just appear like a list of words, their deeper meaning is lost.

### 2.5.5 The Effects of Being an English as a Second Language Learner (ESL)

The literature in this section is focused on word recognition, as this is fundamental to this research work. In considering ESL, the reading process involves the interaction between two language systems. Readers reading in a second language always have access to their first language and often use it as a reading strategy (Carson, Carrell, Silberstein, Kroll, & Kuehn, 1990; Upton & Lee-Thompson, 2001). Indeed, first language (FL) and second language (SL) reading differ in many ways.

#### 2.5.5.1 English: A Complex Language

The identification of regular words has a lower error rate than the identification of exceptional words (Wang & Koda, 2007), as the letter-sound correspondence of exceptional words is not straightforward. Regular words can be easily identified because the words' pronunciation can be generated by some kind of pronunciation rules (Wang & Koda, 2007). It has been claimed further that these pronunciation rules are commonly based on grapheme-phoneme correspondences (GPCs) in the language. For example, the word *must* (a regular word) is always pronounced similarly to *just* and *dust*, but, *have* (an exceptional word) has a different pronunciation to *gave* or *cave*. These critical differences mean that English words are sometimes complex to recognise.

Another example is the words *mint* and *pint*. The pronunciation rules of *mint* (a regular word) cannot be applied to the word *pint* (an exceptional word). Applying the same naming strategies would only lead to error. It has been found that the pronunciations of exceptional words are retrieved from the reader's lexical stores (e.g. Coltheart, 1978; Coltheart, Curtis, Atkins & Hallen, 1993).

#### 2.5.5.2 The Frequency Effect of Written Words

It has been suggested that frequency of words should be a general principle in word identification (Wang & Koda, 2007). For example, Morton (1969) found that readers identify frequently used words more quickly than they do less frequently used words. This definitely had some connection to the regularity of the words, hence the robust finding on the Frequency x Regularity interaction for both alphabetic writing systems such as English (e.g., Hino & Lupther, 1998; Seidenberg,

Waters & Barnes, 1984) and nonalphabetic writing systems such as Chinese (e.g., Fang, Horng & Tzeng, 1986; Hue, 1992; Seidenberg, 1985).

In reading alphabetic languages, spelling-sound regularity has a significant effect on low frequency words, but not on high frequency words (Wang & Koda, 2007). The whole-word pronunciations of high frequency words can be rapidly retrieved from the lexical store before the grapheme-phoneme mapping process is activated (e.g. Coltheart et al., 1993; Coltheart, Rastle, Perry, Langdon & Ziegler, 2001).

Research has also shown the effects of regularity for nonword naming latencies (e.g., Glushko, 1979; Taraban & McClelland, 1987). For example, nonwords derived from regular words (e.g., *nust* from *must*) are named faster than nonwords derived from exceptional words (e.g., *mave* from *have*).

#### **2.5.5.3 The Need to Control the Writing System of SL**

Another important factor that may affect cognitive demand when copying is the writing system used in the copier's FL. Firstly, *different writing systems in the world select different units of spoken language for mapping* (e.g., DeFrancis, 1989; Perfetti, 1999). English, with its alphabetic system, uses phonemes; Japanese kana, with a syllabary system, uses syllables; and Chinese, with a logographic system, uses morphemes or words to represent spoken language. In Chinese, a number of strokes in a logographic may mean a morpheme or a word and hence is also characterised as a morpho-syllabic writing system (Perfetti & Zhang, 1995). This differs to English, which is based on letter-phoneme mapping where letters are grouped together to build up syllables and words. Therefore, an ESL who has Chinese as their FL might encounter difficulty in processing English words and thus need higher cognitive demands.

On top of that, the nonlinear spatial layout of the Chinese writing system differs significantly from the writing of English. Chinese characters consist of interwoven strokes in a square-shaped form, where each stroke has its specified order in forming a word, whereas English uses a linear arrangement of letters. The case is different again in Korean, where although a logographic writing system and square-shaped form are used; the letters are mapped onto phonemes just as in English. Even though this is the case, if a copier's FL uses a Chinese or Korean writing system, this

may have an effect on the cognitive demand they use to process English words, as their logographic is different to the alphabetic system.

The sound-symbol mapping practice in Chinese also differs from the English system. Chinese, specifically, uses semantic radicals (providing meaning information for the whole character) and phonetic radicals (providing sound cues as the syllable level for the whole character), and these mappings are not entirely transparent (e.g., Shu & Anderson, 1997; Shu, Anderson & Wu, 2000).

The regularity of this writing system functions in a different way from the grapheme-phoneme regularity in English, so a Chinese ESL who is not familiar with the grapheme-phoneme may encounter extra difficulties in copying.

Indeed, the studies on the effects of cross-writing-system differences in reading English by learners with a non alphabetic FL background have been investigated by many (e.g., Akamatsu, 1999; Haynes & Carr, 1990; Holm & Dodd, 1996; Jackson, Lu & Ju, 1994; Koda, 1999, 2000; Wang & Geva, 2003; Wang, Koda & Perfetti, 2003). Findings supported the idea that having an alphabetic FL, compared to a logographic one, better facilitates word identification in an alphabetic SL. Muljani, Koda and Moates (1998), who compared Indonesian (an alphabetic system) and Chinese (a non alphabetic system) found that Indonesian ESLs performed more efficiently in an English lexical decision task than Chinese ESLs, as Chinese subjects struggled with the different character form. On the other hand, Wang et. al. (2003) compared Chinese and Korean ESLs in word identification skills and found Korean subjects to be better at identifying words because, as mentioned previously, the Korean logographic writing system and square-shaped form mapped its letters onto phonemes, just as in English. Moreover, they found that Chinese ESLs rely less on phonological information and more on orthographic information when identifying English words, compared to Korean ESLs.

To conclude, an ESL with an alphabetic FL is likely to develop better recognition of English words when phonological decoding is necessary than an ESL with a nonalphabetic FL, such as Chinese or Korean, although Korean learners may do better than Chinese learners. All this discussion on controlling the FL writing system is important, because it stresses the need to choose a suitable participant group in this research work. Because there are effects on the processing costs, it is



important that this research should avoid extra costs, such as the different orthographic system as explained.

### **2.5.6 Summary of Factors that Affect the Cognitive Demand of Copying**

The ability to comprehend words in a sentence impacts upon the understanding of the whole sentence and would therefore affect the length of pauses. For an ESL user, longer pauses might occur because of the complicated process involved in semantic processing for the retrieval of word meanings of unfamiliar words (e.g. Bonin et al., 1997) and unfamiliar spellings (e.g. Bourdin & Fayol, 1994). The higher the familiarity level with words and spellings, the better the comprehension level, so the more competent the ESL user in the language and, hence, the more automatised the retrieval of words in the copying process will be (Alamargot & Chanquoy, 2001). When language or linguistic features become automatised, they are kept in the long-term memory (Lindgren & Sullivan, 2006). With certain features already automatised, there is no further need to use working memory (i.e. short-term memory) to process these behaviours. Processing costs are therefore reduced, which eventually prevents the working memory from becoming overloaded.

Children only achieve the same level of handwriting automaticity as adults at the age of 15 (Sassoon et al., 1986). Prior to this, they would probably struggle when copying and demand more memory space to process both handwriting and language components at the same time. This would then affect pause lengths. The overloading of processing resources would probably be much higher than for adults, due to the simultaneous processing of language comprehension, linguistic features and handwriting production (e.g. the motoric activity). For adults, on the other hand, the handwriting production is more automatic, so pauses are more likely to represent the processing of language components, such as word recognition and language comprehension. Factors described in Section 2.5 are also considered.

## **2.6 Overall Conclusion**

Three main areas which are related to this study have been reviewed: studies of pauses, of sentence-copying and of second language processing. Based on the review, this study recognises that within these areas, there has not been any research investigating pause analysis as a means to measure components of cognitive processes in language processing especially in second language learners, nor is there any approach to using sentence-copying as a method; hence the existence of this exploratory study.

The general aim of the study reported in this thesis is 'to develop a novel method for assessing components of cognitive processes in language processing by analysing the pauses that occur during sentence-copying.' In order to achieve this goal, this study concentrates on free handwriting by examining the competencies of ESL adults.

The review in this chapter directs the researcher's interest to the investigation of pause analysis and sentence-copying, which can potentially be a useful method of language assessment by focusing on the processes that occurs during the test-taking. This method may offer an alternative to assessing language learners through products of knowledge acquisition, and a new perspective in measuring the cognitive processes.

In the exploration of this study, a pilot study and three experiments were devised that focus on three important aspects: (1) the method of copying, (2) suitable sentence stimuli and (3) suitable measures of cognitive processes in language processing overall. As a first move, since this is a novel approach, the pilot study (next chapter) runs as a test to see if such an approach to measure the cognitive processes would be feasible for further exploration(as in Experiments 1, 2 and 3).

## **CHAPTER 3    Pilot Study: The Suitability of Pause Analysis as a Tool for Measuring the Cognitive Processes of Language Processing in Sentence-Copying**

### **3.1 Introduction**

The purpose of this pilot study is to explore the possibility of using pause analysis to measure the cognitive processes in language processing using sentence-copying tasks. A review of the literature (as presented in Chapter 2) showed that pause analysis has not yet been adopted as an approach in measuring language cognitive processes. Therefore, this pilot study will address three fundamental issues that provide the preliminary background support required for such an approach. These are: (1) the methods of copying, (2) the types of sentences to be copied and (3) the measures of cognitive processes in language processing that can be derived from the pauses affected during the process of copying.

The first aim is to explore the methods of copying that could differentiate individuals of varying levels of nature of language processing. This study focuses on two methods that were considered able to distinguish the levels of competence. These two methods are known as Immediate Copying (IC) and Initial Reading (IR).

IC requires participants to copy the sentence stimulus as soon as it is presented; meaning that they would need to access knowledge they already store in their memory instantly. The level of familiarity has a crucial effect on the speed of copying, as it allows participants to retrieve information from memory quickly. Familiarity levels affect the chunking size. For example, one user might be able to group a number of words together, in a bigger chunk size. This is usually the case for a competent language user. But others may only be able to process one word at a time, which is a smaller chunk size. There may also be incidents where participants chunk at syllable level, a smaller chunk size again; this is often the case for a less competent language user. In summary, the level of familiarity determines the chunking process, which influences the pause lengths.

IR, on the other hand, allows participants to read and familiarise themselves with the stimuli for 30 seconds before proceeding to the copying task. Words can be rehearsed, as the allocated time provides the participants with an opportunity to register the words in their short-term memory

(STM) and to recall or retrieve them from their long-term memory (LTM). This interval between reading and copying also allows time for participants to recognise words and even to strategically plan the best way to chunk groups of words. Thus, IR has an overall potential benefit for all participants. On the one hand, a more competent participant may perform better, therefore increasing the amount of differentiation between them and less competent language users. Competent participants might experience a 'ceiling effect', where they reach the maximum level of performance they are capable of. This is an advantage because it could increase the differentiation. However, allowing the 30 seconds would also allow the less competent user to improve their performance.

Both IC and IR are assumed to be able to differentiate levels of competence. However, the effects obtained are different. IR may offer some advantages to participants of all competence levels, but there is no reliable evidence yet that confirms which method is better. This study will therefore explore this further.

The second aim of the pilot is to explore the different types of sentences used. The correct choice of sentence used for copying is also aimed at being able to clearly differentiate levels of competence. For the purpose of preliminary testing to see if this approach is feasible, it would be sufficient to be able to assess general language competence, not taking into account any specific aspects of the language. Therefore, the sentences adopted in this exploration vary across the types and means of manipulations.

'It is a sunny day' is an example of an easy sentence built up from high frequency words. It can easily be recognised by all levels of competence, but it is important to establish whether it can differentiate these levels of competence. Proverbs, such as 'it is raining cats and dogs', may not be recognised by all levels of competence. Even so, could this proverb differentiate the levels of language processing, since the words are of high frequency? On top of that, sentences can also be very difficult to understand when bombastic words are used. This may increase the level of difficulty in the process of comprehension and copying, especially for less competent language users. Given the range of different sentence types, and the lack of prior research in assessing their suitability, this study explores a number of sentences that were considered able to differentiate levels of language processing. These sentences include technical sentences, proverbs, garden paths and simple sentences, and are each further varied by using words of different frequencies,

changing the presentation of the sentence – such as the omitting spaces between words –jumbling words within a sentence, and jumbling letters within a word.

In order to develop a method to measure the cognitive processes of language processing, the study requires an approach that only reflects the nature of language processing, and is not affected by other skills such as handwriting. To achieve this, the study looked into the application of baseline writings, aimed at showing that the measurement of cognitive processes in language processing is not affected by unrelated factors that vary between individuals, such as speed of writing. With this in mind, the baseline comprises the most highly practised letters, such as the letters of the alphabet, and the most frequently written words, such as participants' personal names, to capture basic writing skills. The pauses captured from the copying of baselines are then used to investigate whether there is any relation to language processing, using correlations, which will be carried out in this investigation.

The third aim of this chapter is to find a suitable measure for the cognitive processes of language processing which could be used to calculate the pauses captured and hence be able to evaluate an individual's competence. The measures explored in this chapter are the Median, Mean, Quartile Ratio, Quartile Difference and Normalised Quartile Difference. The pauses, which involve different levels of processes represented by hundreds of milliseconds, are identified as pauses within a letter (L0), pauses between letters in a word (L1), pauses between words (L2) and pauses between phrases or group of words (L3). In order to measure the cognitive processes of language processing, these pauses of L0, L1, L2 and L3 are calculated using a suitable measure. Since it is unclear which the suitable ones are, the pilot will explore all five measures.

As an addition to finding a suitable measure, there is also a need to validate whether the pauses are actually measuring language competence. Participants were therefore ranked according to their language knowledge level, and a correlation was performed between the pause values and the rank order. The validation is performed to get a general picture of whether the method is feasible. It is expected that the result would show a strong correlation if it is able to measure the cognitive processes.

The sentence stimuli design will be explained in the next section.

## 3.2 The Sentence Stimuli Design

This section presents the sentence stimuli design, which consists of the baseline stimuli (Section 3.2.1) and the sentence stimuli (Section 3.2.2). The overall aim of this section is to find a suitable sentence stimulus that will show a clear differentiation between competent and less competent language users. Predictions are made for each stimulus, attempting to foresee their suitability for sentence-copying activity and measuring of cognitive processes. In order to define the levels of competence, predictions are made as to the likely pause lengths and chunking sizes that will occur. Accordingly, the next section will first examine the assumptions for baselines, and then look at the sentences.

### 3.2.1 The Baseline

Baseline stimuli are stimuli that are designed to capture the basic writing skills of participants in an effort to ascertain whether there is a correlation with language processing. It is predicted that baseline copying should not have any relation to language, but instead should be able to correlate well with the sentences because of the fluency of writing the letters of alphabets. *Basic writing skills* are revealed through the writing of highly practised letters and very familiar words, which have minimal processing demands and require limited use of the working memory. Figure 3.1 reveals the stimuli that are considered to represent basic writing skills, which will be explored in this experiment.

1. Alphabet letters in Upper Case.  
*e.g. A B C D E F G H I J K L M N O P Q R S T U V W X Y Z*
2. Alphabet letters in Lower Case.  
*e.g. a b c d e f g h i j k l m n o p q r s t u v w x y z*
3. Jumbled letters in both Upper and Lower case.  
*e.g. m P S z b F g l N Q r A x J o U c e H T w k V I D y*
4. Combination of 2, 3, 4 and 5 letter consonant words.  
*e.g. 'zjbmt lzp bvjh dw qmyjd fkz pvkx vm hbln zdfm'*
5. Name writing in Lower case only, repeated three times.  
*e.g. john doe john doe john doe*
6. Name writing in Upper case only, repeated three times.  
*e.g. JOHN DOE JOHN DOE JOHN DOE*

Figure 3.1 The Baseline Stimuli

Below explains the functions of each stimulus and the predictions that were made as to each one:

- (1) Baselines 1 and 2 aim at capturing pauses in conditions of general fluency of letter writing in upper or lower case. Data collected from these baseline stimuli can be compared with the letter writing data from the sentence stimuli. Copying the letters of the alphabet requires less processing. Assuming that all participants are well-versed in letter writing, there would be a longer pause at the beginning of the letter 'a' or 'A' and shorter pauses between the rest of the letters. This happens because language users do not need to retrieve or recall letters of the alphabet from their LTM, or even store the input in their STM, so the task is performed with a high degree of automaticity. This automaticity allows processing costs during sentence-copying to concentrate on language-related processes. The manipulations applied to baselines 3 and 4 are targeted towards observing the user's fluency in writing the letters of the alphabet when they are jumbled, both in terms of their alphabetical position and the letter-case. The jumbling of upper and lower letter-case tests the fluency of writing in either condition. The random alphabetical position purposefully disrupts the fluency of copying adjacent letters (e.g. abc, mno, pqr) that was measured in baselines 1 and 2.

In baseline 4, the fluency of writing letters in a non-alphabetical position is still being assessed, but the manipulation is additionally targeted towards observing chunking patterns when letters are combined into word-like compounds.

- (2) In summary, the manipulations applied to baselines 3 and 4 are introduced in order to increase the level of unfamiliarity and difficulty in copying the letters of alphabet. Participants would need to concentrate more than with baselines 1 and 2 to avoid making mistakes, thus slowing down the process of copying and increasing pause length. Baselines 5 and 6 constitute the writing of the participants' own name as retrieved from memory. In a literate person's writing career, their personal name is the most common and highly practised word they will write. Therefore, the writing of personal names is likely to be an automatic process, without any processing demands on memory. Participants of all competency levels should be able to write their names fluently, producing short L2 (word) pauses.

To conclude, the researcher predicts that: (1) baselines 1 and 2 will constitute a good indicator of pauses at letter-level (L1); (2) baselines 3 and 4 will provide reliable insights into how the participant chunks at syllable-level (L1); and (3) baselines 5 and 6 will be a good indicator of pauses

at word-level (L2). All in all, it was predicted that participants who are able to write letters and words fluently when given baseline stimuli would be able to copy the sentence stimuli with the same fluency; that there will be a positive correlation between the two. This experiment, however, seeks further to confirm whether performance in copying baseline stimuli has any relation to language processing.

### 3.2.2 The Sentences

In search of a suitable sentence, 9 types of sentences were explored, comprising 37 sentences altogether. These sentences were chosen based on the premise that they are able to discriminate varying levels of competence and thus act as an indicator of language competence. As mentioned in the introduction of this chapter, the sentences comprise technical sentences, proverbs, garden paths and simple sentences, and each of these is further varied by using words of different frequencies, changing the presentation of the sentence – such as by omitting spaces between words – jumbling words within a sentence and jumbling letters within a word. A summary of these sentences are presented in Table 3.1 below.

Sentence Type	Sentence No.		Sentences and Manipulation Description
	IC	IR	
A	7, 8,	26, 27	Technical sentences with specialist terms from specific fields; e.g., ‘... the fuzzy quantization observed in predissociation.’
B	9, 10	28, 29	As group A, but without spaces between words; e.g., ‘Manymutationscandestabiliz...’
C	11, 12	30, 31	Words from a proverb, jumbled; e.g., ‘are than bigger eyes your belly.’
D	13, 14	32	Letters in every word jumbled, except for the end-letters (first and last letters); e.g., ‘Tihs is bcuseae the ...’
E	15, 16, 17	33, 34, 35	Garden path sentences, where the initial meaning is misleading; e.g., ‘The horse raced past the barn fell.’
F	18, 19	36, 37	Common words, with no spaces in between; the last and first letters of successive words are high frequency bigrams; e.g., ‘whathaveshedonenowthatsheshould...’ (i.e., ‘th’, ‘es’, etc. are common pairs).
G	20, 21	38, 39	Two proverbs combined by selecting alternate words; e.g., ‘All actions that speak glitters louder is than not words gold.’
H	22, 23	40, 41	As group G, but with three proverbs combined; e.g., ‘Bad beauty appearances news is are travels only deceptive fast skin-deep.’
I	24, 25	42, 43	As group F, but with lower frequency words; e.g., ‘tanglerequirenaughtidleredictobtuse.’

Table 3.1 The Sentence Stimuli



In Table 3.1, the far left column presents the nine sentence types, labelled from A to I. The next two columns show two sentence numbers, assigned to each of the two methods of copying, IC and IR. Sentence type D has only one sentence assigned to IR, as one of the sentences was excluded because it was not suitable. However, the findings do not show any effects from this limitation, so it was considered acceptable to proceed. In the far right column is the description of the sentences and manipulations, with examples.

Further descriptions of these sentences and the predictions of sentence-copying effects are explained in the following sections. Each of these will conclude with an assessment of the potential of the sentence to measure language competency.

#### *A: Technical sentences with specialist terms from specific fields*

Sentence type A uses sentences from academic journals that use technical terms, such as law, mathematics and engineering. The sentences are grammatically correct, but the technical terms may be difficult for someone who is unfamiliar with the relevant academic field. A competent participant who is also a competent language user, may be familiar with the technical terms and thus be able to chunk groups of words (i.e. at word group processing), producing long L2 pauses. On the other hand, participants who are less familiar with the language, and by extension its technical terms, might find it difficult to comprehend the whole sentence, or may struggle with the technical terms. With these less competent participants, there is a tendency for chunking to occur at word level or even at syllable level. This would result in long L2 and L1 pauses.

In brief, this type of sentence is able to assess participants' competence level by testing the participant's ability to comprehend the whole sentence with regard to the technical terms.

#### *B: Technical sentences with specialist terms, without spaces between words*

Sentence type B is similar to A, but with a heightened difficulty level. This is achieved by omitting the spaces between words, which impedes the reading and recognition of single words, thus increasing the difficulty of comprehending the whole sentence. Competent participants, who are familiar with the technical terms, may still be able to distinguish and thereby chunk the words, even though the omission of spaces might slow down the process of recognition. Less competent participants might not be able to recognise some words, especially the technical terms, and might end up using chunks at Syllable level processing, producing longer L2 or L1 pauses at each chunk.

In brief, this type of sentence, similar to the previous sentence type, is able to assess participants' competence level by testing the participant's comprehension level with regards to the technical terms, but with the added difficulty of distinguishing between words. In addition, processing the whole sentence meaning would definitely be a greater challenge, especially to the less competent participants.

### *C: Words from a proverb, but jumbled*

Sentence type C uses proverbs, but the word positions have been jumbled. The difficulty of recognising the proverbs has been increased because of the jumbling effect. However, competent participants might still recognise the proverb and able to recall it from the list of jumbled words, hence enabling them to process the copying more quickly than less competent participants, using a larger chunk size, at Word group level processing. This is because the less competent participants are less likely to be familiar with the proverbs, and may assume that the sentences are just disconnected word lists, and hence lack the advantage that benefited the competent participants. For the less competent language users, each word in this stimulus will be processed on its own, which accounts for longer L2 pauses.

In brief, this type of sentence is able to assess participants' competence level by testing their familiarity with proverbs, and their ability to connect the words to build the meaning of the whole sentence.

### *D: Letters in every word jumbled, except for the end-letters*

Sentence type D uses normal sentences with correct grammatical structure, but all of the words contain jumbled letters. The letters are randomly jumbled, except for the end-letters, the first and last letters of the word. This manipulation eventually disturbs the spelling of the word and is therefore predicted to interrupt the word recognition processes. Competent participants who are familiar with English words may still be able to recognise the words through their overall shape and the letters contained within the word (Section 2.5.2.2), despite them being jumbled. Copying them, however, is a challenge. Competent participants are likely to experience a conflict between the automaticity of writing the words using their correct spelling and the instruction to copy them jumbled, exactly as presented. This may slow down the process of copying, resulting in word- or syllable-level processing, producing increased overall L2 and L1 pause lengths. This effect could be milder for a less competent participant, who might not recognise the words to begin with and

might even give up trying to comprehend the words and sentence. This is likely to result in their carrying out mainly syllable level processing, producing long L1 pauses from the very beginning.

In brief, this type of sentence is able to assess participants' competence level by testing the participant's ability to recognise words despite the jumbled letters. Since the jumbling is at letter level, it is uncertain whether participants could process the whole sentence meaning. This leaves a question that needs to be further explored.

#### *E: Garden path sentences, where initial meaning is misleading*

Sentence type E uses garden path sentences, whose ambiguity lies towards the end. This ambiguity affects the understanding of the meaning of the entire sentence. The sentence may appear normal, with correct grammar, to all participants, but those who are not familiar with garden paths might pause longer at the ambiguity point. Thus, participants are tested for the ability to semantically process the whole sentence in their search for meaning. Long L2 or L3 pauses are expected to occur at the ambiguous points. The difficulty in retrieving meaning, however, might only affect competent language users, who will be surprised by the change in the expected meaning. It is less likely that less competent participants would notice, as they are not that well versed in such types of sentences. Hence, longer L2 pauses at the ambiguous points are more likely to be observed from the competent participants.

In brief, this type of sentence is able to assess participants' competence level by testing their ability to comprehend the sentence meaning, with regards to the ambiguous point.

#### *F: Common words, with no spaces in between them, where the last and first letters of successive words are high-frequency bigrams*

Sentence type F uses simple sentences with high frequency words. However, the difficulty level in recognising the words is increased firstly by the omission of spaces between words, and even further by the end-letters between words forming high-frequency bigrams. As an example, in 'whathave', the bold letters 'th' are a high-frequency bigram in English, with **t** as part of the word 'what' and **h** as part of the word 'have'. The combination of bigrams aims to make participants confused and hence less able to distinguish the correct word. This might lead them to mistakenly grasp wrong words, which in turn affects the overall meaning of the sentence. For example, the word 'shedone' can be read as 'shed one' or 'she done'. Competent participants might be able to recognise the words more quickly, producing Word group level processing, with long L2 pauses.

However, competent participants are also more prone to making recognition mistakes, as the omission of spaces disrupts the comprehension of the sentence as a whole. Less competent participants may find it difficult to distinguish the words at all, therefore demanding higher processing costs in memory for the search and retrieval of similar familiar words. Less competent users are therefore likely to produce many small syllable level chunks, with long L2 and L1 pauses in between these chunks.

In brief, this type of sentence is able to assess participants' competence level by testing their ability to distinguish high-frequency words, with omitted spaces between words and the use of bigrams at the end-letters.

#### *G: Two proverbs combined by selecting alternate words*

Sentence group G combines two proverbs by alternating their words, producing a jumbled version of the two. This jumbling aims to increase the difficulty of recognising the proverbs. A competent participant, who is familiar with the language and its proverbs, may be able to recall them both from memory. Their jumbled state, however, interrupts the copying process. Language competence still proves instrumental in allowing participants to carry out Word group level processing, with long L2 pauses at the beginning of the sentence. This is due to the competent user's ability to recognise words. However, it is time-consuming for them to attempt to make sense of the jumbled proverbs, and realise that the given sentence is actually composed of two proverbs. A less competent participant might not recognise the proverbs at all and treat the sentence as a disconnected word list. Long L2 pauses are therefore expected to occur, along with Word level processing.

In brief, this type of sentence is able to assess participants' nature of language processing by assessing their ability to grasp, from the disconnected words, that the sentence consists of two proverbs, and thereby testing their familiarity with proverbs.

#### *H: Three proverbs combined by selecting alternate words*

Sentence type H is similar to G, but uses three proverbs instead of two. With type G, competent participants were more likely to recognise the proverbs because of their familiarity with the language; the pattern of jumbling can be recognisable. When three proverbs are used, however, the competent participants would find it significantly more difficult to recognise the proverbs. Thus, the stimulus appears much more like a disconnected word list than when using only two

proverbs. Less competent participants, who do not recognise the proverbs at all, would also be likely to treat the stimulus as a word list just like the competent participants. Hence all levels of competence are likely to carry out word level processing, with many long L2 pauses.

In short, sentence stimuli in groups G and H are more distracting to competent participants than to the less competent.

*I: Lower-frequency words, with no spaces in between, where the end-letters are combinations of high-frequency bigrams*

Sentence type I is similar to F in that spaces between words are omitted and the end-letters between words form high-frequency bigrams. However, type I uses low frequency words taken at random from the dictionary. The selected low frequency words are not likely to be familiar to participants of all competence levels. This is, therefore, predicted to be the most difficult of all sentence stimuli. This part of the experiment tests participants of all competence levels. A competent participant requires more processing from the working memory to store, recall and retrieve unfamiliar words. This activity increases pause lengths at all levels and the pattern of chunks is likely to be inconsistent. It is predicted that the difficulty of distinguishing unfamiliar words will result in Syllable level processing, with some long L1 pauses. Less competent participants, able to distinguish few or none of the low-frequency words, are likely to produce many small syllable level chunks, resulting in many long L1 pauses.

In brief, it is clear that this type of sentence is difficult for all competence levels to process, but is unclear whether it can differentiate them. The sentence type does increase the demands of processing, meaning that every participant will be tested to their maximum knowledge capability.

### 3.3 Method

Following the predictions of the stimuli, this section will present the methodology used for the pilot.

#### 3.3.1 Participants

Five participants took part in this study. They ranged in competence from native to non-native speakers of English. An assessment of their knowledge of the English language was obtained through interview sessions. They were ranked in the order of the English fluency observed from

the interview session, taking into account biographic knowledge, such as country of origin, years of experience in an English-speaking country, and education level. These participants were labelled as A, E, S, U and T, with A being the most competent and T the least competent. Table 3.2 summarises the participants' background following the order of competence.

Initial	Background
A	PhD Student; British; Female; 25 years of English experience; English native speaker.
E	Master's Student; Dutch; Female; 23 years of English experience; English as a Second Language.
S	Master's Student; Malaysian; Lived in Norway for the past 10 years; Female; 35 years of English experience; English as a Second Language; English Teacher.
U	PhD Student; Malaysian; Female; 22 years of English experience; English as a Second Language.
T	Bachelor's Student; Malaysian; Male; 23 years of English experience; English as a Second Language.

Table 3.2 Participants' Backgrounds

### 3.3.2 Apparatus

All writing activities were performed on a standard graphics tablet (Wacom, Intuous3) with an inking pen on a piece of paper taped to the tablet. The paper was printed with a grid of 20 x 13 cells or boxes, each measuring 1cm x 1cm. Participants were trained to write one letter in each box and to omit spaces between words in the sentences. The use of these boxes is to ensure that a distinct pause is associated with each and every letter produced. This would not be the case for cursive writing.

This study adopts a similar method to that used in previous experiments, conducted in lab conditions, using sentence-like stimuli (e.g. Cheng & Rojas-Anaya, 2006 & 2007). A custom made program, TRACE (Cheng & Rojas-Anaya, 2004), was used to record all writing actions and to extract the data on pen positions and timings. The analysis of the pauses was performed manually in Microsoft Excel.

### 3.3.3 Procedure

The participants were asked to copy the sentences given to them at their own normal writing speed. Sentences were visible one at a time, and each sentence was visible until the copying was completed. Each trial began with the six baseline stimuli. This allowed participants to become

familiar with the activity, the tools used and the new environment. Three of the participants did the copying under the IC condition first, where they were required to begin copying as soon as the researcher said 'start', while the other two started under the IR condition, where a maximum of 30 seconds were allowed for an initial reading before the copying began. Given that the copying task involved 37 sentences, it could be considered quite lengthy and time-consuming. The participants were tired towards the end, but the results are nevertheless clear, indicating that fatigue does not seem to have been a factor systematically interfering with their performance. The participants were also trained to begin each sentence with a 'hash' (#) in order to ensure that the process of writing is well underway before the first stimulus letter was copied. Overall, the five participants performed 43 different writing tasks, producing a total of 215 sentences.

### 3.3.4 Overall Structure of this Experiment: The Variables

Figure 3.2 introduces the overall structure of this experiment, giving an overview of the main points being taken into consideration as part of the effort in achieving the overall aim of this thesis (as represented in Level A).

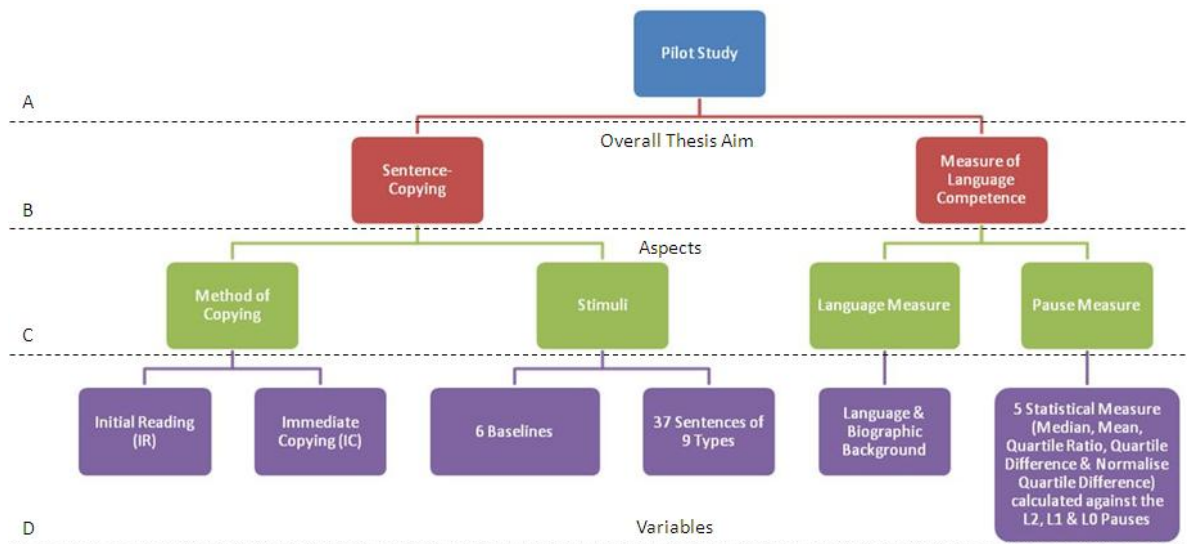


Figure 3.2 The overall structure of this experiment

Four aspects presented at level C, group together the core components to this study (level D). These variables are discussed in this chapter, in the investigation of the suitability of pause analysis to measure language in general. In investigating the most suitable method of copying, the study compares IR and IC. In search of suitable sentence stimuli for sentence-copying, it explores 37 sentences, grouped into 9 types of sentences, as well as 6 baseline tasks. In the assessment of a

suitable measure, 5 different statistical measures are explored, which are validated by being correlated against the language and biographic backgrounds.

### 3.3.5 Analysis

All copying data captured by TRACE were pre-analysed to extract and calculate the pauses between each stroke. This pause data were then mapped onto the letters of the words in the sentence manually, using Microsoft Excel, to distinguish the pause levels (i.e. L2, L1, and L0). This process was carried out for all 43 stimuli of each individual participant. These pause levels were then calculated with each of the five statistical measures chosen, and then compared with the participant's competence ranking gathered from the biographic background, in order to measure their language competence.

## 3.4 Results

The results will be presented by first looking into the suitable method of copying and the measure of cognitive processes. These two separate factors were analysed at the same time. Pauses were measured as part of the attempt to find a suitable copying method, and different measures were used to calculate these pauses, in the attempt to find a suitable measure, thus both aspects were explored concurrently. Following this, the result presents the findings on the distribution of pauses at L0, L1 and L2 pause level, aggregated by sentence type, which allows the observation of the pattern of pause distribution. Finally, the result presents the observation made for the baselines, in the attempt to confirm that the sentence-copying activities are only measuring the cognitive processes, and not other, unrelated, factors. Hence, correlations were conducted between the baseline pauses and the competency rank order, and between the baseline pauses and the sentence-copying pauses.

### 3.4.1 Suitable Method of Copying Sentences & Measures of Cognitive Processes in Language Processing

The pauses captured are used as the basis to explore the measures of cognitive processes in language processing: the Median, Mean, Quartile Ratio ( $Q3/Q1$ ), Quartile Difference ( $Q3-Q1$ ) and Normalise Quartile Difference ( $(Q3-Q1)/Q1$ ). At the same time, this will help show which methods of copying would be suitable for this study. To assist the analysis, the sentences are grouped



according to their sentence type. Hence, there will be nine data points (group A–I), representing the nine types of sentences.

Table 3.3 reports the Pearson correlation (1-tail) values calculated for both the IC and IR experimental conditions for the overall pause values (consists of all L0, L1 and L2) across all five measures. The Median has the strongest correlation with language competence in the IC condition. A detailed analysis conducted on the IC sentences found a consistent strong correlation with the Median, where the lowest correlation among the sentence types was  $r(3)=0.82$  ( $p<.05$ ). The Mean also exhibited strong correlations, even when there were a number of non-significant correlation values among the sentences. Detailed analysis carried out under the IR condition showed more variable correlation values across all sentence stimuli types, as shown by the average IR correlation value in Table 3.3. Summing up, the Median measure of all the pause values (L0, L1 and L2) in the IC condition appears to be the most consistent and appropriate choice for measuring the cognitive processes of language processing. Hence, it can be concluded that the Median of all pauses (L0, L1 and L2) may be an effective measure to the nature of language processing, and that IC may be a suitable method for sentence-copying.

Pearson Correlation	Median	Mean	Quartile Ratio (Q3/Q1)	Quartile Difference (Q1-Q3)	Normalised Quartile Difference (Q1-Q3)/Q1
Average for IC	0.95**	0.93*	0.23	0.61	0.14
Average for IR	0.79	0.81*	0.33	0.44	0.01

Table 3.3 Correlation values calculated using Overall Pause Values between the Methods of Copying and the Competency Rank Order. (\* $p<.05$ ; \*\* $p<.01$ ; \*\*\* $p<.001$ )

### 3.4.2 The Distribution of Pauses in Different Types of Sentences

The analysis so far has been based on the overall pause values. This section will now focus on the IC method (found to have a stronger correlation than IR) and the Median measure of segregated pause levels: L0, L1 and L2. Table 3.4 presents a summary of these findings. L1 exhibits stronger significant correlations across all the different types of sentence stimuli, compared to L2 (word) level. L0 provides the weakest average of the correlations mean.

		Pause Levels		
		L0	L1	L2
Type of Sentences	A	0.195	0.970**	0.835*
	B	0.031	0.974**	0.750
	C	0.443	0.878*	0.653
	D	0.754	0.852*	0.705
	E	0.762	0.988**	0.715
	F	0.536	0.962**	0.717
	G	0.567	0.930*	0.771
	H	0.513	0.977**	0.613
	I	0.645	0.954**	0.691
	MEAN	0.494	0.943**	0.717

Table 3.3 Correlation values calculated for each pause level and type of sentence (\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ )

The pattern of pauses was observed at each pause level by looking at the effects of correlation between the pauses of each level (L0, L1 and L2), grouped by the sentence type, against the participant's language performance rank order. The data are presented on a graph, with the rank order displayed on the x-axis, arranged from most competent (left) to least competent (right). The y-axis displays the pauses in milliseconds (ms). A positive correlation graph is expected, as can be seen from the correlation values displayed in Table 3.4 above. The positive correlation explains that a competent person would be able to perform the copying in a shorter overall time than the less competent person, as the latter would take a longer time to process each of the sentences, because of the lack of familiarity.

Thus, each of the graphs below will focus on the results at each pause level.

Figure 3.3 provides a close-up view of the results focusing on L0 (stroke) pauses. These show a small pause range, of 70ms to 160ms, compared to other pause levels across all five participants. The correlation values of L0 across all sentence groups and participants vary and are not significantly strong (refer to Table 3.4). It is expected that a strong significant correlation would preserve the participant's rank order; that positive correlation values will be seen, decreasing from the most competent (A) to the least competent (T). Observing the graph in Figure 3.3 demonstrates that the correlations do not preserve the rank order, consistent with the weak correlations found as in Table 3.4.

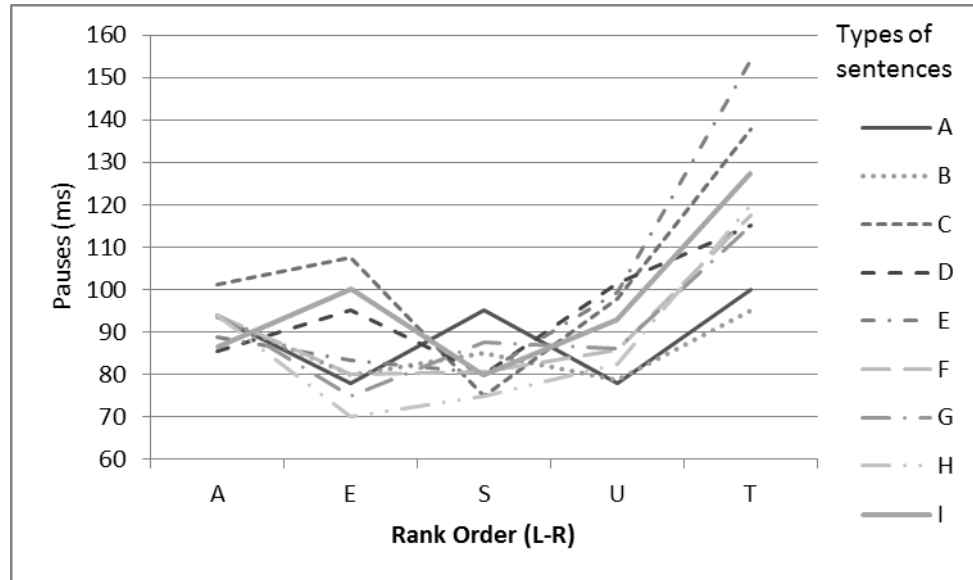


Figure 3.3 A close-up view of L0 pause levels

Figure 3.4 provides a close-up view of the results focusing on L1 pauses. These show a wider pause range, of 140ms to 440ms, than in L0. The correlation values of L1 across all sentence groups and participants are consistent and significantly strong (refer to Table 3.4). It is expected that a strong significant correlation would preserve the participant's rank order; that positive correlation values will be seen, decreasing from the most competent (A) to the least competent (T). Observing the graph in Figure 3.4 demonstrates that the correlations do preserve the rank order, consistent with the strong correlations found in Table 3.4, with an average of  $r(3)=0.943$ ,  $p<.01$ .

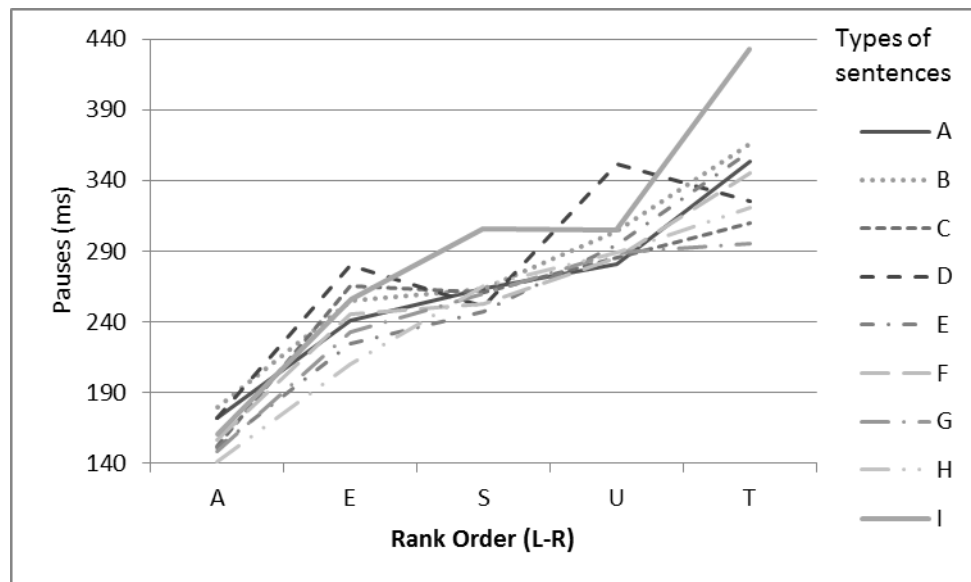


Figure 3.4 A close-up view of the L1 pause levels, demonstrating the predicted preserved rank order

Figure 3.5 provides a close-up view of results focusing on L2 (word) pauses. The L2 pauses show the widest pause range, 200ms to 1100ms. However, the correlation values of L2 across all sentence groups and participants are not consistent and the correlation is therefore not significantly strong (refer to Table 3.4). The average of the correlations across all nine types of sentences was  $r(3)=0.717$ . It is expected that a strong significant correlation would preserve the participant's rank order; that positive correlation values will be seen, increasing from the most competent (A) to the least competent (T). Observing the graph in Figure 3.5 demonstrates that the correlations do not preserve the rank order, consistent with the weak correlations found as in Table 3.4. The correlations for participant S and U, in particular, show the most variance across the sentence types.

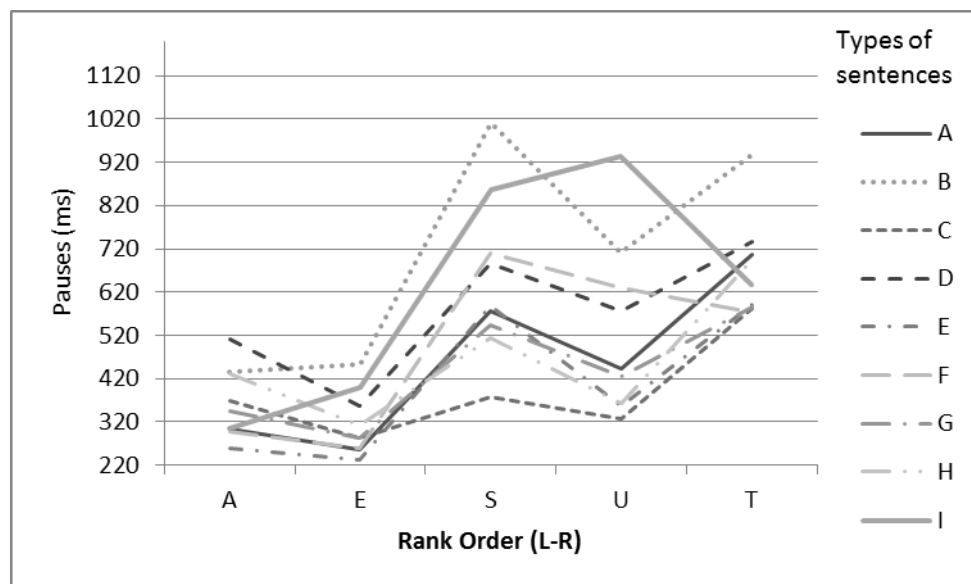


Figure 3.5 A close-up view of L2 pause levels

As an overall view of the pause distributions across the baseline writing, as shown in Figure 3.6, an increase of pause distribution can be seen from L0 to L2.

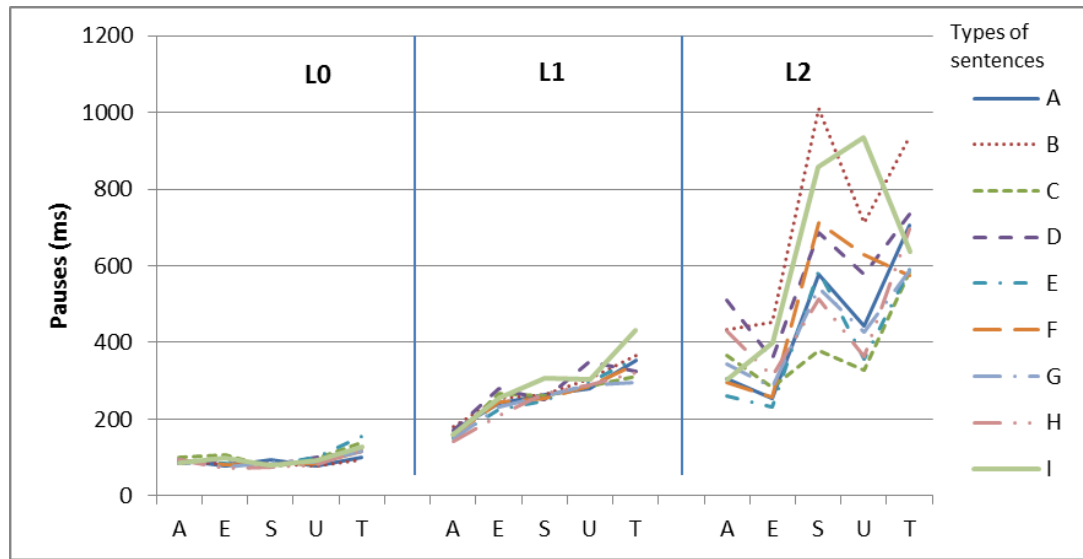


Figure 3.6 An overview of Pause Distribution at L0, L1 and L2 pause levels across all participants and sentence types

### 3.4.3 The Baselines

The results in this section present the correlation between the baseline tasks and the language rank order, and between the baseline tasks and sentence copying. It was predicted that there would be no relation between the baseline and language performance, because baselines are measuring basic writing skills which have reached the level of automaticity. However, it is predicted that the baseline would correlate greatly with the sentence stimuli, because basic writing skills are at the core of both tasks. The pauses are based on the Median of L1s, following the previous findings.

#### 3.4.3.1 Baseline Versus Language Competence

Figure 3.7 reports the data obtained from the baseline copying, focusing only on the Median of L1 pauses because the findings obtained from the sentence-copying shows a strong correlation effect with Median of L1s. The result demonstrates that the baseline stimulus of *lower case name writing* has the strongest correlation value,  $r(3)=0.958$  ( $p<.01$ ), and is also able to preserve the predicted rank order, thus suggesting that this is the appropriate baseline. The graph shows a wide range of pauses, from 150 to 700ms, for L1, although the largest variance can be seen with participant T.

Other baseline stimuli were not able to preserve the rank order, although they still have strong correlations; *alphabet letters in upper case* achieving  $r(3)=0.829$  ( $p<.05$ ), *jumbled letters with varying upper and lower cases* achieving  $r(3)=0.875$  ( $p<.05$ ), and *combination of 2, 3, 4, and 5 letter consonants* achieving  $r(3)=0.863$  ( $p<.05$ ). The other two baselines showed weak correlations:

*alphabet letters in lower case* yielded a significantly low value,  $r(3)=0.551$  and *upper case name writing* yielded a value of  $r(3)=0.659$ .

Overall, the results obtained from the baselines suggest that there is some relation between the basic writing skills copying and language competence, which is surprising. Further argument for this will be presented in the Discussion section (3.5).

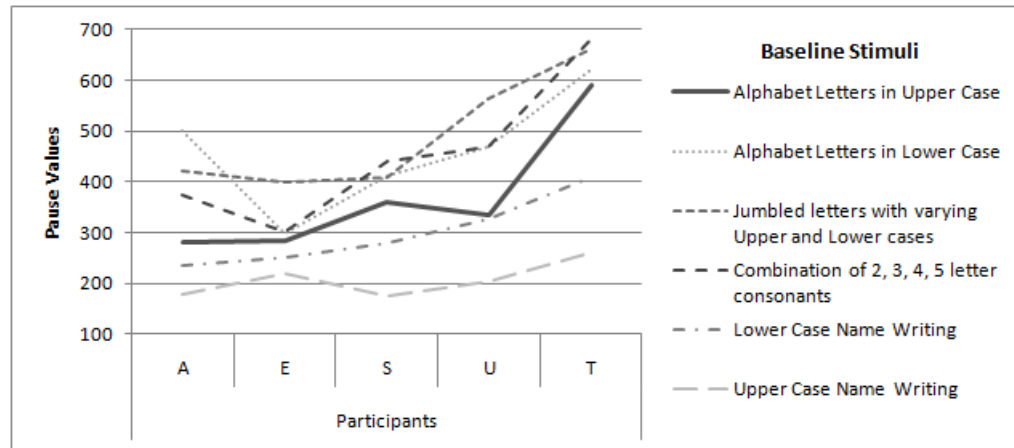


Figure 3.7 Baseline L1 (Median) Pauses correlated against the Language Competence Rank Order

#### 3.4.3.2 Baseline Versus Sentence-copying

Figure 3.8 presents the results of the baseline data correlated against the sentence stimuli. However, the baseline data were only taken from the lower case name writing, as this was shown to provide the most promising data, with the strongest correlation values across all the six baseline types. The sentence stimuli data, on the other hand, were taken from the median of L1 and L2 pauses. The two lines on the graph shown represent the correlation values across all sentence stimuli, with the solid line representing the median of L1s and the dotted line representing the median of L2s.

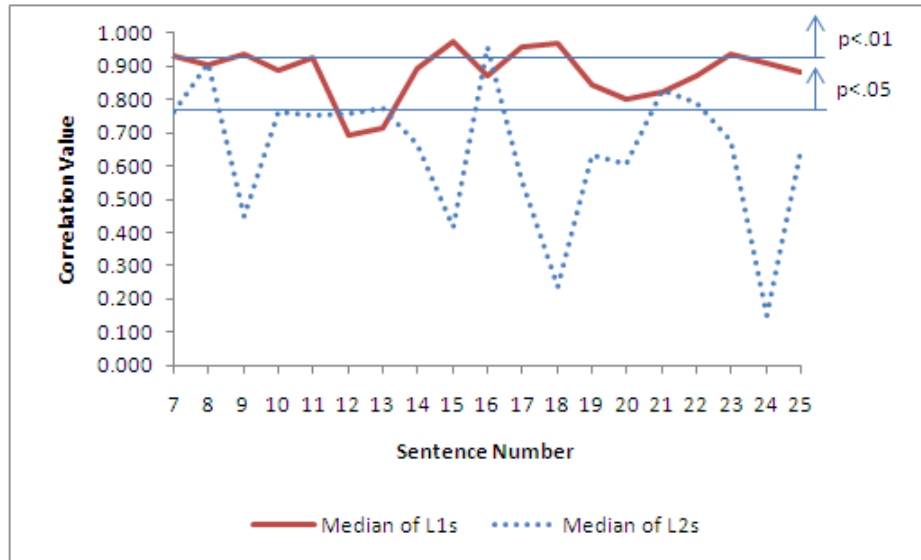


Figure 3.8 Correlation values based on median of L1 and L2 pauses between the baselines of lower case name writing against the sentence stimuli

The results gained from this analysis were interesting, as they were able to show strong and almost consistent correlations of the median of L1s across all the sentence stimuli except for sentence numbers 12 and 13. Despite these exceptions, the mean of the overall correlation value for the median of L1s is still strong;  $r(3)=0.881$  ( $p<.05$ ).

The correlation results between median of L2s and the sentence stimuli are represented by the dotted line. The results here, however, do not provide a strong overall (mean) correlation;  $r(3)=0.650$ . Detail observation on the correlation values across all the sentence stimuli was also not consistent.

### 3.5 Discussion

The aim of this pilot experiment was to explore the suitability of pause analysis in measuring the cognitive processes of language processing through sentence-copying. More specifically, the pilot explores (1) methods of copying, (2) suitable sentence stimuli and (3) suitable measures of cognitive processes. It was predicted that pause analysis, with different pause levels (i.e. L0, L1, L2, L3) captured for the different levels of processing (Level 1, 2, 3, 4), would be able to measure the cognitive processes underlying language processing. Furthermore, it was predicted that participants with a good language background would be able to perform the sentence-copying smoothly, hence producing shorter pause lengths, whereas a less competent participant with a

poor language background would be able to copy less smoothly and therefore produce longer pause lengths.

In accordance with the prediction made, the pauses captured agree with the rank order of participants, producing a positive correlation. In brief, a competent participant with a strong language knowledge background would be able to produce shorter overall pause lengths than a less competent participant, because of their differing abilities to comprehend the language. The pilot experiment has demonstrated that the technique of sentence-copying is relevant, as it shows that there is a temporal signal in the lower-level cognitive processes of writing (i.e. stroke, letter and word) captured through the activity of copying that can be related to the measuring of cognitive processes in language processing in general.

Further to the findings above, the pilot has been able to find a measure, the Median of pauses (L0, L1 and L2), appropriate for this study, and has discovered that IC is a suitable method for copying. L1 pauses unexpectedly show significant correlations in this study; it was expected that L2 (word) level pauses would be more revealing, as letters do not portray any language meaning. More discussion on this is presented in Section 3.5.2. It was also surprising to see that baselines, which consist of basic writing skills, could correlate well with language competence (Section 3.5.4). It was also discovered that the length of pauses increases from L0 to L1 and L2 in a consistent manner ( $L0 < L1 < L2$ ), as shown in Figure 3.6. This indicates the increase in the amount of processing required prior to the copying of successively bigger chunk sizes (Cheng & Rojas-Anaya, 2006, 2007 & 2008). The strong correlations obtained from the analysis showed that all the sentence types tested can be used for measuring the cognitive processes, but were not able to distinguish the most suitable type of sentence for this purpose.

A detailed discussion of each of these findings is presented in the next few sections.

### **3.5.1 What is the Suitable Method for Copying: IC or IR?**

IC was found to be the most suitable method for copying, as it is able to differentiate between varying levels of competence in a more controlled manner than IR. Correlation values across all sentences against the rank order of competence were found to be more consistent with IC than with IR. This is reflected by the average of correlations of IC, with  $r(3)=0.95$  ( $p<.01$ ) for Median and  $r(3)=0.93$  ( $p<.05$ ) for Mean (Table 3.3). This was further confirmed by the strong correlations



obtained at the more segregated pause levels of L0, L1 and L2, as presented in Table 3.4. On top of that, the ability to preserve the rank order, as strongly shown with the L1 pauses of IC (Figure 3.4), supports the choice of IC as the suitable method of copying for this study.

IC and IR differ by the initial time of 30 seconds given to prepare. IR, which allows participants to prepare beforehand, gives the opportunity for all participants of different levels of competence to apply individual strategies in order to be able to copy fluently. For example, somebody who is competent might pre-plan how to best copy the sentence, calculating the points at which to stop and the number of words to capture at each point of copying, which helps to make copying very smooth. The extra time given initially also gives the opportunity to comprehend the sentence meaning before starting the activity of copying, which would then decrease the need to comprehend while copying. Less competent participants will also be able to use this initial time to prepare a copying strategy, especially when the sentence is considered difficult. Thirty seconds is considered sufficient time to familiarise oneself with the word shape and the spellings of unfamiliar words. All of these IR effects differ from IC, because IC does not give the opportunity to prepare, instead forcing participants to use all the knowledge they have on the spot. IC allows the measuring of cognitive processes in language processing based on the current knowledge capability.

The results show that the extra time to prepare allows copying strategies to be developed, thereby introducing many new processes. IR therefore increases the difficulty of measuring the cognitive processes as it does not provide significant results that represent the pure processing of language comprehension. IC, however, forces the comprehension of language to be done during the copying activities, which means participants are being assessed purely based on the knowledge that they have at that moment in time. Not allowing them to prepare beforehand helps to isolate the processes that need to be focused on for the purpose of measuring the cognitive processes in language processing.

Therefore, as a conclusion, choosing IC as the suitable method of copying is also reflected by the strong correlations gained from the experiment.

### 3.5.2 What is the Effective Measure for Cognitive Processes in Language Processing?

Median was found to be a reliable measure and hence may be an effective measure of cognitive processes in language processing. Even though the results for IC from the Mean were similar to the Median (Table 3.3), a thorough examination of the correlation values against each sentence and sentence group reveals an inconsistency of correlation values with the Mean. The Median, however, has yielded the most reliable results, with consistent correlation values across all sentence types, especially for L1 pauses. In addition, the Median was able to reflect the rank order of the participants' competence levels.

In terms of pause levels, the findings showed a strong correlation with L1 pauses. L0 and L2 pauses were found to be insignificant. Logically, it was thought that L2 pauses would be more meaningful than L1 pauses and should therefore, provide a better correlation for language competence. The results show, however, that this is not the case. One possible reason for this may be the different types of sentences employed in the experiment; as they are not consistent, this may have affected the employment of different processes. On top of that, the range of participants' competence levels and their biographic backgrounds were too wide, which might have affected the way each individual processed these sentences. Further, the inconsistency of the kind of manipulation used on the sentences may have affected the processing at word level, is reflected by the poor correlation with L2s. The effects on the L1 pauses may be because of the chunking processes occurring at the syllable level, and thus, the consistent correlation results with L1. Overall, the results suggest that more focus should be placed on L2, in order to be able to test at word level, which may be more meaningful in language terms. This will be considered in the next chapter.

Figures 3.9 and 3.10 show an example of L1 pause distribution for the most competent participant (A) and the least competent participant (T) in this experiment, respectively.

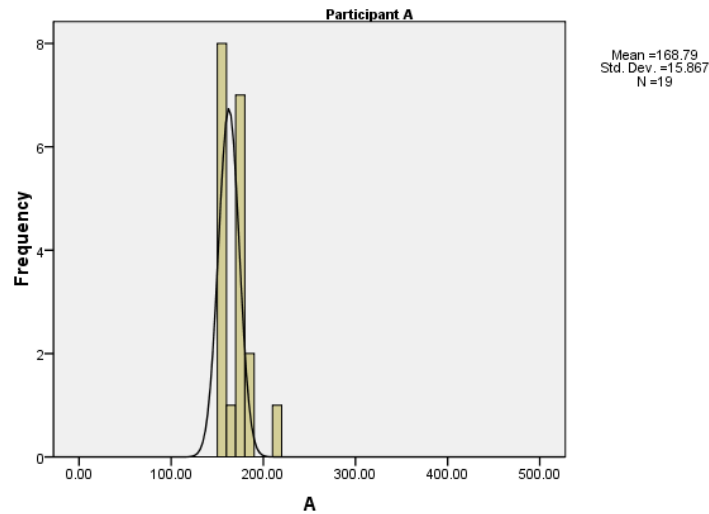


Figure 3.9 An example of L1 pause distribution for the most competent participant (A) in this experiment

In Figure 3.9, participant A shows a greater focus on the lower pause levels, ranging from 150 to 250ms, indicating that they are fluent in copying letters and familiar with the language. A higher level of competency requires less processing to copy familiar words and letters. However, the case is different for the least competent participant, T.

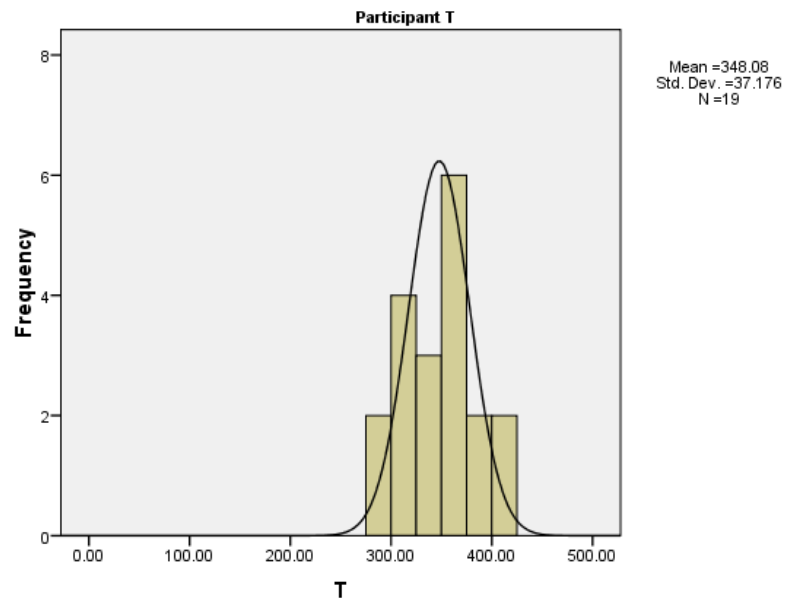


Figure 3.10 An example of L1 pause distribution for the least competent participant (T) in this experiment

Figure 3.10 shows a much higher pause range, between 250-450ms. T takes a longer time to process the copying, reflected by the letter level copying. This clearly differs from participant A.

This is likely to be because of the level of difficulty with and the unfamiliarity of the language. There is also the possibility that T may be dyslexic, which may need to be considered.

As a conclusion, this pilot was able to show that the median of pauses is a good measure. However, the result from the pause levels is not what was expected. It is surprising that L1 shows a better correlation, although this may be affected by the various manipulations employed in the sentences. This suggests a need for further investigations in confirming the levels of pauses.

### **3.5.3 What is the Appropriate Sentence Type for the Purposes of this Thesis?**

The strong significant correlations found across all sentence types under the IC condition, as presented in Table 3.4, indicate that all the sentences are suitable for measuring the cognitive processes. Even though the limited number of participants with a wide language background restricts in-depth analysis of the effective sentence stimulus, strong significant correlations with IC and overall pauses that are consistent across all sentences indicate that they are all potentially able to differentiate between the different levels of competence. Figures 3.4 and 3.6 illustrate the preserved predicted rank order represented by the sentence stimuli. Although the rank order is only preserved for L1 pauses, there is likelihood that it would also be preserved for L2 pauses if the limitations were removed. This suggests the need for further explorations in search of a suitable sentence type, focusing more on the L2 (word) level pauses.

### **3.5.4 What is the Suitable Baseline Stimulus?**

Some of the baselines have yielded significant correlations with language competence, with 'lower case name writing' having the most significant value,  $r(3)=0.958$  ( $p<.01$ ). The results have shown that basic writing tasks, such as copying the letters of the alphabet and name writing, are good measures of cognitive processes in language processing based on L1 pauses. This is unexpected, as it does not make logical sense that writing names or letters of the alphabet would be able to measure a person's nature of language processing. It could be that the choices of participants recruited for this study is too broad and therefore not suitable to establish an appropriate baseline. Or it may be that participant T is dyslexic, thus affecting the pause data. The evidence for this possibility is that T produces long pauses across all sentence stimuli, even when copying at letter level. However, the exact reasons are still unclear, suggesting a need for further investigation.

The correlation between the strongest baseline (lower case name writing) and the sentences provides this study with a consistent strong correlation across all the sentences except for sentence numbers 12 and 13, which indicates that the prediction that the sentence-copying consists of proficient letter writing is true. This means that the writing of letters in the sentence-copying shows the same pause patterns as the writing of letters in the baseline. Sentence numbers 12 and 13 produced a weak correlation, which may be because of the nature of the sentence difficulty. Thus, it can be concluded that baseline copying may be able to confirm whether the proficiency of letter writing in the sentences is maintained through the correlation results.

### 3.6 Conclusion

In this experiment, the method of copying, the measure of cognitive processes in language processing, the most suitable sentence type for measuring the cognitive processes and the baselines were all considered. IC was identified as a suitable method of copying, median as the suitable statistical measure to calculate the pauses and L1 pauses as the most promising pause level to consider. However, the present data have not been able to distinguish between the sentence stimuli to identify the most effective sentence to differentiate the levels of competence. As discussed, it may be that the number of participants was too small to enable further comparison to be made. However, this is a pilot study, an attempt to establish which ideas could be worth exploring. It may be premature at this stage to determine the sentences that would be effective for such a study, which suggests the need for further investigation in the next experiment (Chapter 5).

On the other hand, the baseline tasks (in the case of lower case name writing) were found to be a good indicator of the letter copying ability found within the sentence copying (in the case of lower case name writing). However, the findings that show a good correlation between the baseline and language competence rank order might still need further exploration.

It can be inferred from this pilot study that it is feasible to apply pause analysis in the study of copying English sentences for assessing general language processing. Further, the pilot study also generates questions such as (1) what affects whether pause lengths are long or short?(2) what are the underlying cognitive processes occurring during the activity of copying? (3) how do participants process the words in the sentences? and (4) what kind of sentences are suitable for

measuring the cognitive processes of language processing? Before moving on to the next experiment (Chapter 5), Chapter 4 will present the Theoretical Model of Copying, a theoretical assumption developed in order to understand how the process of copying takes place.

## CHAPTER 4 A Theoretical Model of Copying (MoC)

### 4.1 Introduction

This chapter will present the *theoretical model of copying (MoC)*, a model developed in order to assist in (1) analysing and describing the flow of sentence-copying processes, (2) identifying the factors that might affect the length of pauses amongst participants of varying competence levels, and (3) designing sentence stimuli.

The MoC comprises several processes predicted to be present during the activity of copying sentences. Although the model was established during the implementation of Experiment 2 (Chapter 6), this chapter is presented earlier in the thesis, as it helps in the understanding of the processes that occur during the experiments that are introduced from Chapter 5 onwards.

The rationale behind the model is that the lower-level processes that exist in the activity of copying are very complex and so cannot be easily explained. As explained in the previous chapters, pauses are assumed to represent cognitive processes, which this thesis hypothesises to be related to language knowledge. Pause lengths signify the demand of *processing capacity* needed in the working memory. The use of *processing capacity* of an individual is affected by their level of familiarity and competence with the language, wherein these will then affect the pause lengths. For example, a competent person who is familiar with the language might demand less processing capacity and thus produce shorter pause lengths. In contrast, a less competent person who is not familiar with the language will demand more processing capacity (perhaps even maximising the usage of working memory spaces), hence producing longer pauses. Using the MoC, this study hopes to explain these occurrences during the activity of sentence-copying, in order to achieve the major aim of this thesis: “to develop a novel method for assessing components of language competence by analysing the pauses that occur during the activity of sentence-copying.”

The proposed MoC was inspired by earlier models: the *Dual Route Cascaded Model* (Coltheart et al., 2001), the *Extended DRC Model* (Verhoeven et al., 2006), the *Language-Processing Model* (Patterson & Shewell, 1987) and the *Handwriting Production Model* (Van Galen, 1991).

The term ‘copying’ includes the reading process, built on word recognition, and handwriting production. As copying is partly related to the graphic transcription process (i.e. handwriting), the model adopts the lower order modules of Van Galen’s Handwriting Production Model (1991)

including systems for: spelling, the selection of allograph, size control, and muscular adjustment. Although participants in this thesis were assumed to possess automatic handwriting processes, the model has been designed in such a way as to cater for a wider population, and thus may be suitable for future nature of language processing research.

Research on visual word recognition has advanced considerably in the past decade and has provided a general framework for the study of word reading in English (Patterson & Shewell, 1987; Van Orden & Goldinger, 1994; Plaut et al., 1996; Coltheart et al., 2001) and in Dutch (Verhoeven et al., 2006). There are also studies that relates to spelling factors (Kandel et al., 2006a, 2006b, 2009; Lambert et al., 2008) in which they looked into the influences of syllable such as graphemes and phonemes. Early work on word recognition and reading models was instigated by research into reading difficulties, such as dyslexia. These models were based on reading aloud and therefore focused on speech output and processes involving the phonological system ('phonological output lexicon'), which do not relate to the activity of copying. However, these phonologically driven processes are not related to the copying tasks, instead, it has been found that the act of copying is clearly phonologically driven (Kandel et al, 2006a, 2006b, and 2009). There remains the possibility, however, of a degree of 'sub-vocal' involvement.

In addition to this line of research, Rayner and Pollatsek (1989) studied eye movement control processes in order to try and detect when reading processes decelerate, which can be construed as an indication of poor comprehension.

On the higher level of language testing research, a recent study by Weir and Khalifa (2008) looked into the cognitive processes involved in the reading section of the IELTS test. Their work produced a model designed to understand the processes involved in the reading section. Their research did not involve the development of methods through which to measure reading processes. However, the lower section of their model, which focuses on word recognition, lexical access and comprehension, is useful for the model.

Models of skilled reading (Scarborough, 2001) and of reading fluency (Hudson et al., 2009) have also been developed in the area of literacy. These two methods proposed the same lower level components of language, word recognition and language comprehension, but with different purposes. Scarborough emphasises that his model is geared towards achieving a fluent coordination of word reading and comprehension. Hudson et al., on the other hand, discuss that



the elements of reading fluency examined are restricted to the processes that are important for accurate and rapid word recognition. Although these two models are influential, their focus is towards understanding how best to achieve reading competence, which is not the intention of this study.

In conclusion, there have been models developed by a number of researchers from different fields in relation to language, in the attempt to understand word recognition, reading performances, and skilled reading, amongst other things, but these have not been extended to measuring the cognitive processes underlying language processing in ESLs. The MoC therefore aims to promote the understanding of the processes involved during the activity of copying, as well as to enable the design of sentence stimuli by providing simulations of sentences to confirm their suitability. It also provides a framework to support data interpretation.

The following section will first introduce the MoC by describing it in detail, and then show some general simulation examples as an overview of how to apply the MoC.

## 4.2 Background

Important information is needed to serve as a guide to understand the model and its workings. This basic knowledge: the levels of processing and the characteristics of words used in the study are briefly explained below.

### 4.2.1 The Levels of Processing

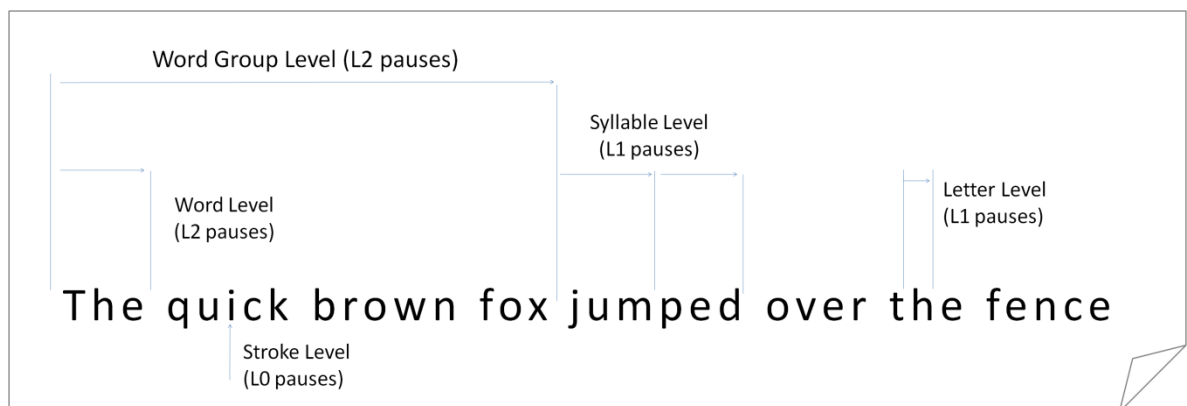


Figure 4.1 Illustration of the pause levels (in brackets) and the levels of processing

Figure 4.1 shows a similar illustration to Figure 1.1 but with additional explanations with regards to the levels of processing involved in this study. The figure explains part of a sentence with examples

of pause levels and the levels of processing positioned within the particular sentence that is captured during the process of copying. For example, a competent participant is expected to have many long L2 pauses because they are chunking groups of words (i.e. *word group level processing*). For someone copying word by word (i.e. *word level processing*), long L2 pauses are expected between the words, in this case, between “The”, “quick”, “brown”, and “fox”, for example. The L2 pauses will be longer than the pauses between letters in a word (L1) or pauses within a letter (L0). L1 pauses occur successively within a word, for example, between the letters “t”, “h” and “e”; i.e. *letter level processing*. There are also the possibilities of *syllable level processing* with L1 pauses, especially when the particular words are too long, take for example the word ‘jumped’, where there the word is chunked as ‘jum’ and ‘ped’. Finally, L0 pauses occur for letters that have more than one stroke, such as the two strokes in the letter “t”. These pause levels and levels of processing are important, as they determine the values of the pauses captured and the position of its occurrences, which are vital to the analysis of this study. Further illustration on how pause levels relates to the levels of processing is shown in Figure 4.2.

Levels of Processing	Pause Levels	Example of Chunking
Word Group Level	L2 pauses	The quick brown fox   jumped over the fence.
Word Level	L2 pauses	The   quick   brown   fox   jumped   over   ...
Syllable Level	L1 pauses	...   qui   ck   bro   wn   fox   jum   ped   ...
Letter Level	L1 pauses	T   h   e   q   u   i   c   k   b   r   o   w   n   ...
Stroke Level	L0 pauses	l   7   h   e   q   u   i   .   c   l   /   \   b   ...

**Figure.4.2 Example of chunking in relation to the levels of processing and pause levels**

Levels of processing differ from pause levels, in that the former represents the position of the chunking processes determined by participants, whereas pause levels represent the value of the pauses that occur at specific positions in the sentence-copying process. It is critical to define this earlier on, as these terms will be used throughout this thesis. Unlike pause levels, where the position of pauses is definite, levels of processing vary between participants, according to individual’s levels of competence. For example, a competent participant might have many chunks of word groups, while a less competent participant might have chunks at syllable level. Figure 4.2 therefore shows examples of chunking processes (represented by the ‘|’ (border) symbol), with the levels of processing mapped to the pause level values. In the process of copying, the chunking at *word group level processing* and *word level processing* is associated with L2 pauses, while the

chunking processes that involve syllables and individual letters are associated with L1 pauses. Stroke level processing is associated with the L0 pauses.

Figure 4.2 differs to Figure 1.1 with the additional information on (1) *word group level processing* which does not refer to L3 (pauses between phrases) as the pause values and (2) the introduction of the *syllable level processing*. The *L3 pause* is a general term for phrases however is not applicable to use here in the analysis because of the fact that it is difficult to point the exact position of a phrase chunking. Different people have different phrase chunking sizes; therefore, it was not possible to simply say that certain particular position is a phrase chunk for that particular individual. However, it was possible to refer to the long pauses that occur at word, letter and stroke level, because of its obvious positions. Therefore it would be wiser to refer to the long pauses that occur after a group of words as L2 pauses. As for the *syllable level processing*, the same reasoning to the phrase chunking applies. It is difficult to specify exactly where a syllable level processing would occur. However, the long pauses that occur within a word as L1 pauses can be pointed out.

#### 4.2.2 Word Characteristics

The sentences used in this study contain words that are categorised into three groups according to their characteristics: *length*, *familiarity* level and *frequency* level. These three characteristics are specifically defined to assist in sentence design and also to assist in understanding the copying processes in the MoC. They assist sentence design in that they enable predictions to be made of where in the sentence one would stop to copy, as well as the possible pause value, depending on the word characteristic encountered. Defining word familiarity and frequency, on the other hand, depends on the classification and assumptions made in regard to which words are considered easy and which difficult. At times, the MoC also assumes that the participant is familiar with certain words or syllables through their first language. In Experiment 2 (Chapter 6), high frequency words were defined as words within the first 1000 words in the British National Corpora, and low frequency words as words at 1001 and beyond.

Classifying words accordingly assists this study in understanding where long or short pauses might occur and why long pauses occur at that particular position. For example, a long pause might occur because of a low frequency word, a long word, or a word that is unfamiliar to that individual. There may be other reasons, such as that the copier is unfamiliar with the spelling, or because it

seems like they recognise the word, but the letters are jumbled. When a word contains jumbled letters, a competent person would still be able to recognise the word because of the letters in it, but this would be more difficult for those who are less competent. To those who are less competent, the jumbled letters may seem like a group of letters put together at random.

On the other hand, short pauses may occur because the copier is familiar with the word, the word is of high frequency, or the word length is short. However, short pauses may also occur in individuals confronted with long words, unfamiliar words or low frequency words, because they grouped letters in order to be able to complete copying. As a whole, there is no straightforward interpretation of what causes long and short pauses, and there are many assumptions as to what causes the pause length. This assumption of word characteristics is made to assist this research work, and is not based on previous research.

The three characteristics (length, frequency and familiarity) are assumed to be related to each other. For example, a short word such as 'bug' is a *high frequency word* and a *short word*, and will have a greater chance of being *familiar* to an individual compared to a 'bog'. To conclude, when words are of higher frequency and short in length, it provides a greater familiarity to an individual. The same word may also fall under the category of unfamiliar for a particular individual, if one is not familiar with the language, especially those who have minimal exposure to it. Short words would probably carry a higher probability of being chunked as a whole word, or even as part of a group of words. However, long words have a higher chance of being chunked into two separate chunks, especially when the words are unfamiliar and of low frequency.

Table 4.1 shows a summary of the different combinations of word characteristics.

	SHORT		LONG	
	Familiar	Unfamiliar	Familiar	Unfamiliar
High Frequency	Sho-Hi-Fam	Sho-Hi-Unfam	Lo-Hi-Fam	Lo-Hi-Unfam
Low Frequency	Sho-Low-Fam	Sho-Low-Unfam	Lo-Low-Fam	Lo-Low-Unfam

Table.4.1 The labelling used to describe the word characteristics

The word characteristics are defined based on the assumptions below:

- Words can be short or long; the short having 5 letters or fewer.

Words are considered short when they have 5 letters or fewer. They are considered long when they have more than 5 letters. The length is particularly crucial when they are unfamiliar and are low frequency to the participant.

- The familiarity level is divided into two categories: familiar and unfamiliar words.

Familiarity is a proxy for the nature of language processing in sentence-copying. It is assumed that words can also be categorised as familiar and unfamiliar, which would depend on the level of frequency. Here, this study suggests that words amongst the first 1000 in the list of Vocabulary Profile – British National Corpora (Nation, 2007) will be known as familiar words, and the rest will be known as unfamiliar words.

- Word frequencies are also grouped into two categories: high and low.

Similarly here, high frequency words belong to the first 1000 words in the Vocabulary Profile – British National Corpora (Nation, 2007). Low frequency words refer to those ranking at 1001 and below in the vocabulary profile. The labels used in Table 4.1 represent the combinations of these word characteristics, as well as showing their rank in terms of difficulty of word recognition, *Sho-Hi-Fam* being the easiest and *Lo-Low-Unfam* being the most difficult. As an example, the word ‘there’ would be classified as Sho-Hi-Fam or ‘short, high frequency, and familiar’ for both competent and less competent users. The word ‘vivid’ is also short, but could either be familiar or unfamiliar depending on the competence level of the user, and is a low frequency word, according to the BNC word list (see Section 6.2.4, Laufer & Nation, 1995). A competent person might therefore treat the word ‘vivid’ as Sho-Low-Fam: ‘short, low frequency and familiar’, while a less competent one would treat it as Sho-Low-Unfam: ‘short, low frequency and unfamiliar’.

In the case of comprehension (refer to Section 2.5.4), it can be seen that there are two levels of semantic processing in order to build up understanding of the sentence. There are two levels of processing information: at individual word level, familiar English syllables make up a word, and at sentence level, these familiar words then make up an understandable sentence. The understanding of how comprehension is built up is as straightforward as (1) with no comprehension, the sentence would appear like an arbitrary list of words, but (2) with comprehension, the sequence of words is determined. *Neighbouring words* assist in building up the overall meaning, and therefore make copying much easier. Neighbouring words are words that appear side by side; knowing the meaning of one word can help in understanding the meaning of the next word and, as a whole, assist in building up the meaning of the whole sentence.

Although less competent individuals with less exposure to English may be able to copy fluently at syllable level, close analysis of the pause distribution may be able to reveal that the copying is similar to copying an arbitrary list of syllables or words. Similar pause analysis on competent individuals with good exposure to English would be able to show clear chunking, with clear measures of pause lengths. However, in some cases where the sentence to be copied is an easy sentence with high frequency words, a flat pause distribution can be seen from a competent individual, because of the ability to copy while reading. This is related to the working memory capacity, which is able to hold information and process it without greater demands, because the words are familiar to the participants.

However, in the case of a less competent participant being given the same easy sentence, with high frequency words, it is not surprising if a clear pause distribution is seen at word level. This pause distribution may show that they are able to process the meaning of the sentence. As a conclusion, given any sentence ranging from easy to difficult, pause analysis may be able to provide data and distinguish competence levels, because of the different processing involved, the impact of word characteristics on the processing, and the different approaches used to comprehend the sentence meaning.

In the case of proverbs as a sentence, similar semantic processing is assumed to be applied. Individual word level processing and whole sentence processing are the possible processes used, especially by Second Language (SL) participants. Proverbs are greatly related to familiarity; if someone is not familiar with a common proverb used, they may not be able to understand it. Even if someone is exposed to the language every day, this does not guarantee that they will know the

proverbs, unless they use it in their daily life. However, again, the reason such a sentence is used is to test whether SL participants process it meaningfully, treat it as a normal sentence, or maybe even process it at word level only. Either way, the distribution of pauses on such sentence type could show interesting findings.

### 4.3 Model Structure

The MoC is presented in Figure 4.3. The input is a *printed text* and the output is a *written text*. The grey boxes represent the processes. The lighter grey boxes represent the internal processes within that particular process. The box with the dashed line represents the syllable chunking process, within which the internal processes are represented by orange boxes. The explanation for each of these boxes is given in Section 4.3.1. The MoC predicts seven different routes to processing the sentence-copying, which are explained through a flowchart in Section 4.3.2.

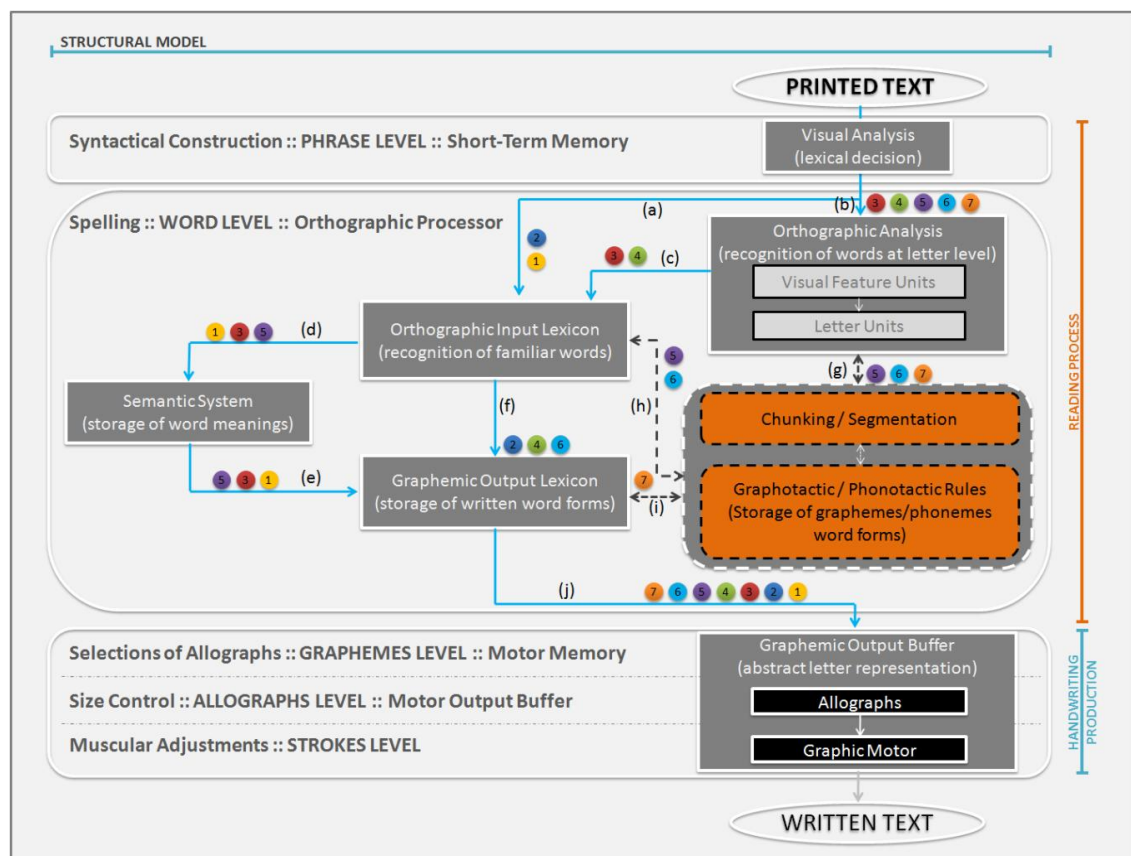


Figure 4.3 Theoretical MoC (Model of Copying)

These routes are demonstrated by the solid and dashed arrow lines, representing the direct route and the parsing route, respectively. The direct routes, also known as lexical route, is where processing is meaningful, while the parsing route, also known as the non-lexical route, is where processing is less meaningful. Detailed guidance for using the model is given in Section 4.3.3, which also shows the *processing load prediction* (PLP), an assumed processing value for the simulations of sentence stimuli. The simulation is performed to test and predict whether a given sentence stimulus is hard or easy to process by people with different levels of competence.

#### 4.3.1 The Components of the Model

This section introduces the definition of each process, Table 4.2, applied in the model in Figure 4.3. A detailed description of some of these processes can be found in Coltheart et al. (2001); this section focuses only on the functions related to copying. In broad terms, the model consists of two major sections: ‘Reading Process’ and ‘Handwriting Production’. The ‘Reading Process’ examines word recognition at four different levels of: word group, word, syllable and letter. These levels of processing are explained in Section 4.2.1. The MoC also introduces two types of routes: the ‘direct route’, for processing familiar words, and the ‘parsing route’, for processing unfamiliar, rare, complex, difficult and irregularly spelt words. For the purpose of making predictions about the processes that take place, a ‘processing load prediction’ (PLP) is used to measure the predicted length of pauses (see Section 4.3.3). Each of the processes that appear in the model is explained in the table below.

Process	Function
1. Visual Analysis	The lexical decision determines the chunking size of words, which depends on the level of familiarity. For example, a competent participant would be able to recognise words by their ‘word shape’ and letters that belongs to particular words and hence is able to chunk a group of words together, as one big chunk. A less competent participant might be able to recognise single words, and therefore have many smaller chunks at individual word level.
2. Orthographic Input Lexicon	This process receives input in the form of groups of words (i.e. phrases) and individual words. It processes the words based on familiarity level. This process makes decisions: if a word is unfamiliar, the process will be referred to the Semantic System, but, if a word is familiar, the process will direct the word to the Graphemic Output Lexicon.
3. Orthographic	This process analyses a word in more detail, examining it at letter level. The Visual



Analysis	<p>Feature Units help to recognise the shape of words (see Coltheart et al., 2001), and then further analyse the Letter Unit. The purpose is to confirm whether there is a need to group letters together as a chunk, within a word.</p> <p>If letters can be taken as a whole word, the process will be directed towards the Orthographic Input Lexicon. If letters need to be grouped by syllable, the process will be directed to Chunking/Segmentation.</p>
4. Semantic System	<p>This process contains an individual database of words which stores word meanings, similar to a dictionary. The Semantic System is part of the Long Term Memory (LTM). This process requires some time in order to locate a word. The process is only used when it is needed in order to assign meanings to words. Once the process is complete, it will be directed to the Graphemic Output Lexicon.</p>
5. Graphemic Output Lexicon	<p>This process stores the written word forms in the Short Term Memory (STM). It also contains the automaticity instructions for writing familiar words and letters. This process will move on to the Graphemic Output Buffer, ready to be written down.</p>
6. Chunking / Segmentation	<p>This process is necessary when copying unfamiliar, rare, complex, difficult and irregularly spelt words. The process of chunking depends on the Graphotactic/Phonotactic Rules. Once the chunking has been carried out, the process will move on to either the Orthographic Input Lexicon or the Graphemic Output Lexicon.</p>
7. Graphotactic / Phonotactic Rules	<p>This process stores grapheme and phoneme knowledge.</p> <p>Participants who are familiar with graphemes will chunk syllables based on grapheme knowledge, e.g. for 'f' → 'ph' as in <i>phantom</i> and 'gh' as in <i>laugh</i>.</p> <p>Participants who are familiar with phonemes will use the knowledge to chunk the letters. There might be the possibility of the 'sub-vocal' approach in chunking at syllable level, influenced by the 'vowel harmony'. As pointed out by Smith (1971) and Kandel et al. (2006), however, there is no need for a phonological process in this model where chunking at syllable level is concerned, especially during the course of copying.</p>
8. Graphemic Output Buffer	<p>This process represents the handwriting production activity. It stores information on a temporary basis and will produce written text as the output. The process also involves the selection of allograph for writing and motor activity for the writing production.</p>

Table 4.2 The functions of each process in the model

#### 4.3.2 The Flowchart

This section presents the flowchart simulation of the model. The letters in the circles are used to explain the possible routes in processing the copying. The flow begins with reading at sentence level, with Lexical Decision determining the size of chunking (group of words or word level). These chunks will be kept in the Short-Term Memory (STM). After deciding on the level of chunking, the process then involves the 'Orthographic Processor', which further analyses the words. However, it is also important to highlight that the input '(a)' which involves groups of words is processed as a group of words, will then go through either '(d)' or '(f)' as one whole word. As a group of words, each individual word is not treated as a single word, therefore does not act as an input to '(b)'. In addition, the choices of routes depend on the competence level of the participants. Finally, 'Handwriting Production' completes the copying task by producing the written output.

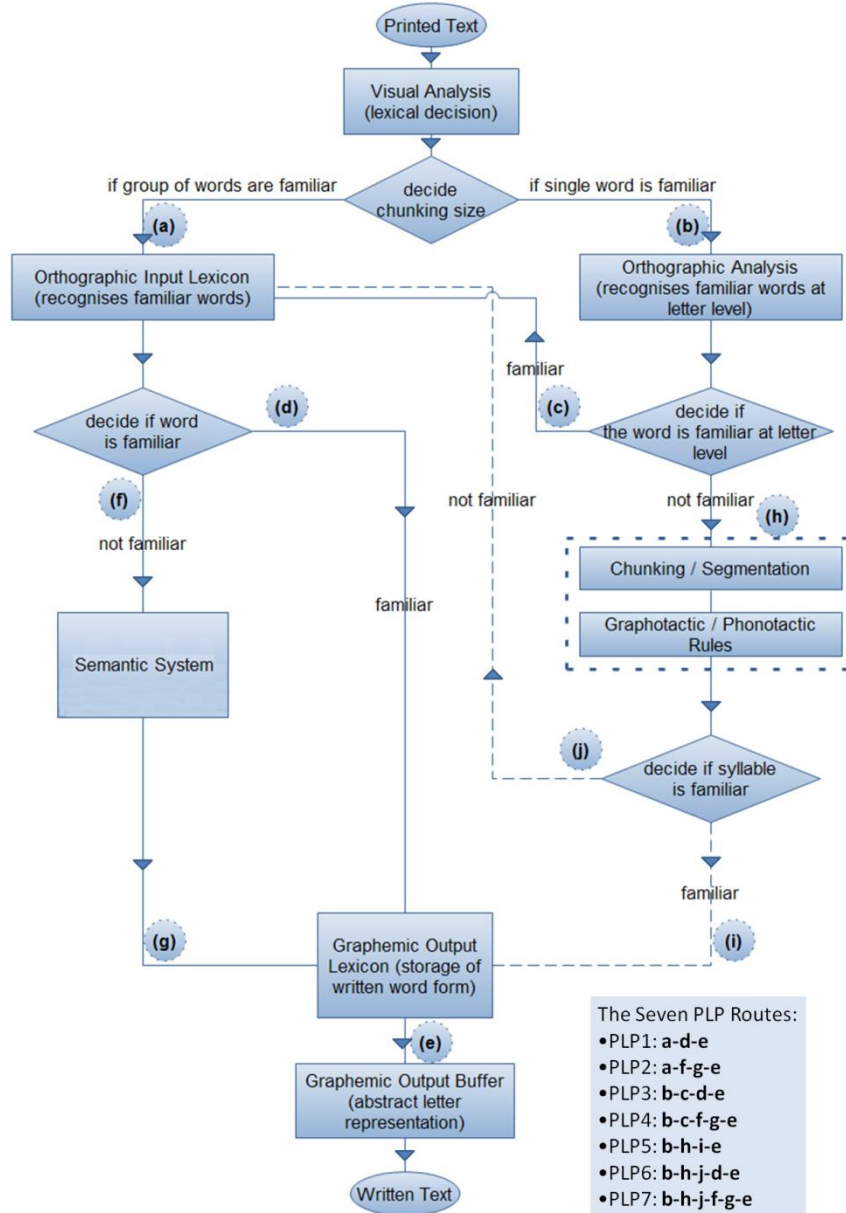


Figure 4.4 The flowchart explaining the decision making processes involved in sentence-copying

Each of these PLPs is explained in detail in the next section.

#### 4.3.3 The Routes & Processing Load Prediction

The model provides seven different routes, each containing a number of processes, which entail different 'processing costs'. A parameter value is set for each route, assuming that each process requires a given amount of time. Further research (e.g. Coltheart et al., 2001) is needed to identify precise parameter values for each process, and this lies beyond the scope of this thesis. However, for the purpose of designing sentences as stimuli and for predicting the pause patterns that will be

produced by more competent (MC) or less competent (LC) users, the nominal parameter value of '1' is assigned to each process, except for the Orthographic Analysis and the processes that involve chunking at syllable level (the dashed box). In the Orthographic Analysis, the value of 0.5 is assigned to both the Visual Feature Units and the Letter Units because they are dependent on each other and hence share the same process. However, the processes that involve chunking at syllable level, containing 'Chunking/Segmentation' and 'Graphotactic/Phonotactic Rules', are assigned a parameter value of '1' each, as they are stand-alone processes. When added together, these parameter values will produce a value for each route, called the 'processing load prediction' or 'PLP'. The flowchart shown in Figure 9 shows these possible routes, represented by letters in circles. The routes are explained in more detail in Table 4.3 below. Further examples of how the model applies to sentence simulation are given in Section 4.4 onwards.

PLPs assist in predicting the overall pause lengths duration of each route. This researcher predicted that graphs of PLP values will show a similar patterns to that obtained from the real copying activities (see Section 4.4 for an example). An MC might use the shortest route for copying the most familiar words or phrases, hence producing shorter pause lengths. However, the route they use might change when words have been manipulated, for example by using irregular spellings, and so they produce longer pauses. By contrast, an LC might use one particular route repeatedly in order to complete copying a single word, or for each word, until the phrase is complete, which would then produce even longer pauses.

Level:	Route:	PLP name & Functionality	PLP Value
PHRASE	PLP1 a-d-e	<p>LPnSR</p> <p>Lexical Phrase non-Semantic Route</p> <ul style="list-style-type: none"> <li>At Phrase level, this is used when phrase chunks contain familiar words.</li> <li>E.g. “it is a fine day” or “I am fine”; taken as one big chunk.</li> </ul>	4
	PLP2 a-f-g-e	<p>LPSR</p> <p>Lexical Phrase Semantic Route</p> <ul style="list-style-type: none"> <li>At Phrase level, this is used when phrase chunks contain unfamiliar words.</li> <li>E.g. “many spectators”; the word ‘spectators’ might be difficult, hence the need to refer to the Semantic System.</li> </ul>	5
WORD	PLP3 b-c-d-e	<p>LnSR</p> <p>Lexical non-Semantic Route</p> <ul style="list-style-type: none"> <li>At Word level, this is used when words are familiar.</li> <li>E.g. the word “library” for a competent person</li> </ul>	5
	PLP4 b-c-f-g-e	<p>LSR</p> <p>Lexical Semantic Route</p> <ul style="list-style-type: none"> <li>At Word level, this is used when words are unfamiliar.</li> <li>E.g. the word “library” for a less competent person</li> </ul>	6
SYLLABLE	PLP5 b-h-i-e	<p>GPnLR</p> <p>Grapheme-Phoneme non-Lexical Route</p> <ul style="list-style-type: none"> <li>This is used on remaining syllables after the initial syllable chunking.</li> <li>E.g. in the word “afternoon”, the copying of syllable ‘noon’</li> </ul>	6
	PLP6 b-h-j-d-e	<p>GPLnSR</p> <p>Grapheme-Phoneme Lexical non-Semantic Route</p> <ul style="list-style-type: none"> <li>This is used when syllable chunking is used for a familiar word; usually it is assigned to the initial syllable chunking of a word.</li> <li>E.g. in the sentence “tanglerequirenaught..”, it is difficult to distinguish words, hence the possibility of syllable chunking of maybe ‘tang’ + ‘ler’ + ‘equi’ + ‘ren’... which are familiar to the participant</li> </ul>	7
	PLP7 b-h-j-f-g-e	<p>GPLSR</p> <p>Grapheme-Phoneme Lexical Semantic Route</p> <ul style="list-style-type: none"> <li>This is used when syllable chunking is used for an unfamiliar word; usually it is assigned to the initial syllable chunking of a word.</li> <li>E.g. in the sentence “tanglerequirenaught..”, it is difficult to distinguish words, and maybe performing syllable chunking is also not familiar to the participant such as ‘tan’ + ‘gle’ + ‘req’ + ‘uir’...</li> </ul>	8

Table 4.3 The Processing Load Prediction (PLP)

#### 4.3.4 Route Predictions

This study has labelled the word characteristic combinations (Section 4.3.2) and the routes (Section 4.3.3) with their predicted values. Table 4.4 matches the word characteristic combinations with the possible routes that could be used to process the copying of each word. This is done irrespective of the impact of competence levels. The potential processing routes are described further below.

Route (PLP) Word Characteristics	Words chunked at Phrase Level		Words chunked at Word Level		Words chunked at Syllable/Letter Level		
	1	2	3	4	5	6	7
Sho-Hi-Fam	✓		✓				
Sho-Hi-Unfam		✓		✓			
Sho-Low-Fam			✓				
Sho-Low-Unfam				✓			
Lo-Hi-Fam	✓		✓				
Lo-Hi-Unfam				✓			✓
Lo-Low-Fam			✓				
Lo-Low-Unfam					✓	✓	✓

Table 4.4 The possibilities of each word characteristic combination to be processed at specific chunking level

At group level chunking process, where familiar words can be grouped together in one big chunk, three words characteristic combinations were identified that allow these words to be copied as a chunk:

1. Sho-Hi-Fam: When words are 'short, high frequency and familiar,' the participant – whether an MC or and LC, may use the PLP1: LPnSR (4), without the need to refer to the Semantic System. e.g.: "it is a fine day".
2. Sho-Hi-Unfam: When words are 'short, high frequency and unfamiliar,' an LC participant may use the PLP2: LPSR (5), and refer to the Semantic System, because of the unfamiliar words. e.g.: "it is a major word".
3. Lo-Hi-Fam: When words are 'long, high frequency and familiar,' a participant – most likely to be an MC – might use the PLP1: LPnSR (4), without the need to refer to the Semantic System. e.g.: "it could be anything".

At word level chunking process, seven words characteristic combinations were identified that allow these words to be copied as individual words:

1. Sho-Hi-Fam: When words are 'short, high frequency and familiar,' a participant might use the PLP3: LnSR (5), without the need to refer to the Semantic System. This is possible for both the MC and LC, but is more likely for the LC. e.g. "could"; "come"; "good".
2. Sho-Hi-Unfam: When words are 'short, high frequency and unfamiliar,' a participant – most likely the LC – might use the PLP4: LSR (6) and refer to the Semantic System. e.g. "blow"; "store"; "lord".
3. Sho-Low-Fam: When words are 'short, low frequency and familiar,' a participant might use the PLP3: LnSR (5), without the need to refer to the Semantic System. This is possible for both the MC and the LC, but is more likely from the MC. e.g. "virus"; "wreck"; "blown".
4. Sho-Low-Unfam: When words are 'short, low frequency and unfamiliar,' the participant might use the PLP4: LSR (6) and refer to the Semantic System. This is a possible route for both the MC and the LC, but is a more likely one for the LC. e.g. "vice"; "oscar"; "havoc".
5. Lo-Hi-Fam: When words are 'long, high frequency and familiar,' the participant – normally the MC – might use the PLP3: LPnSR (5), without the need to refer to the Semantic System. e.g. "something"; "because"; "anything".
6. Lo-Hi-Unfam: When words are 'long, high frequency and unfamiliar,' the participant – normally the LC – might use the PLP4: LSR (6), but refer to the Semantic System. e.g. "thought"; "probably"; "excuse".
7. Lo-Low-Fam: When words are 'long, low frequency and familiar,' the participant – normally the MC – might use the PLP3: LnSR (5), without the need to refer to the Semantic System. e.g. "hopeful"; "playground"; "cutest".

At syllable level chunking process, three words characteristic combinations were identified that allow these words to be copied by groups of letters. Within this context, PLP6 and PLP7 are expected to occur at the beginning of a word, while the PLP5 takes place within a word at syllable level. Further descriptions are given below:

1. Lo-Hi-Unfam: When words are 'long, high frequency and unfamiliar,' the participant – normally an LC – might use the PLP7: GPLSR (8) for the initial letter of the word, and refer to the Semantic System. They also might use PLP5: GPnLR (6) for processing the syllables within the word. e.g. "intermediate": (8) in (6) ter (6) me (6) dia (6) te

2. Lo-Low-Fam: When words are 'long, low frequency and familiar,' the participant might use the PLP6: GPLnSR (7), without the need to refer to the Semantic System. They also might use PLP5: GPnLR (6) for processing the syllables within the word. This would normally be the case for the MC, especially when faced with irregularly spelt words, like words with jumbled letters. e.g. "dangerous": (8) dan (6) ger (6) ous
3. Lo-Low-Unfam: When words are 'long, low frequency and unfamiliar,' the participant might use:
  - a. PLP5: GPnLR (6) for processing the syllables within a word – for both the MC and LC.
  - b. PLP6: GPLnSR (7), without the need to refer to the Semantic System. This would be the case for the LC.
  - c. PLP7: GPLSR (8), and refer to the Semantic System. This would be the case for the LC.
 e.g. "eavesdropping": (8) eav (6) es (7) drop (6) ping

At letter level, the model predicts that the writing of each letter also constitutes a cost, hence the parameter value of '1' is assigned to each individual letter. For example, with the word "eavesdropping": (8) e (1) a (1) v (6) e (1) s (7) d (1) r (1) o (1) p (6) p (1) i (1) n (1) g. Even though individual letters might not represent a meaningful language unit (with the exception of some letters, such as 'a' and 'l'), they can still indicate handwriting production fluency. As such, they should have a parameter value of their own. Therefore, it is assumed that '1' would be a fair assignment.

#### 4.4 General Simulation of the Model

As a general simulation to demonstrate the flow of processes in the copying of English sentences through the model, a brief description is provided here based on two imaginary extreme examples of a most competent (MC) and a least competent (LC) participant. First, this will look at different levels of processing, and then will demonstrate the simulation at sentence level. The example below demonstrates the processing that might occur for three different types of chunks:

- Word group level chunking
- Word level chunking
- Syllable level chunking



Given the conditions used in the experiments, it is assumed that word group level chunking would be the largest possible chunking level participants could carry out and that syllable level chunking is the most basic chunking size that is meaningful in terms of linguistic features. A sentence example is therefore given for each condition listed above, to show the kinds of processes that might occur at each chunk.

#### 4.4.1 Word Group Level Chunking

Given the sentence 'are than bigger eyes your belly', where the words in a proverb have been jumbled, it is assumed that the extreme examples of both the MC and LC are able to chunk at group level. The number of words in a chunk, however, differs between the competence levels. This is illustrated in Figure 4.5. The letter combinations in the arrows represent the route taken in order to process the copying of each chunk. Refer to Figure 4.3 for the route flow.

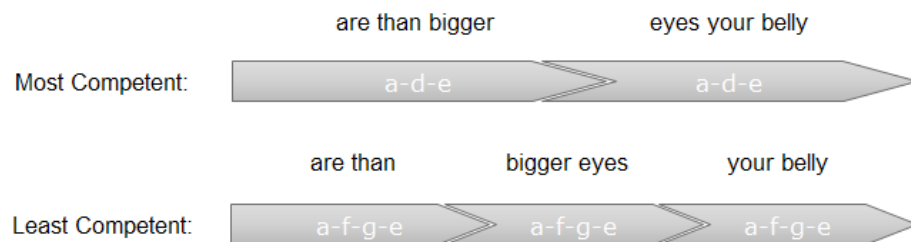


Figure.4.5 Example of chunking at group of words level and the processes that might take place during the sentence-copying

#### 4.4.2 Word Level Chunking

Given the sentence 'tanglerequirenaughtidler', where low frequency words are used and spaces between words are omitted, it is assumed that the MC is able to chunk at individual word level as illustrated in Figure 4.6, but that the LC would probably need to carry out syllable chunking, as demonstrated in Section 4.4.3.

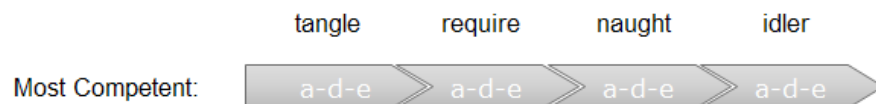


Figure 4.6 Example of chunking at individual word level and the processes that might take place during the sentence-copying

### 4.4.3 Syllable Level Chunking

Given the same sentence, ‘tanglerequirenaughtidler’, the LC would not be able to recognise individual words, as low frequency words are used, and the omission of spaces between words makes it more difficult to identify them. Chunking might instead occur at syllable level, as illustrated in Figure 4.7. Three different routes might be followed, depending on the level of familiarity of each syllable. A combination of different routes could also occur.

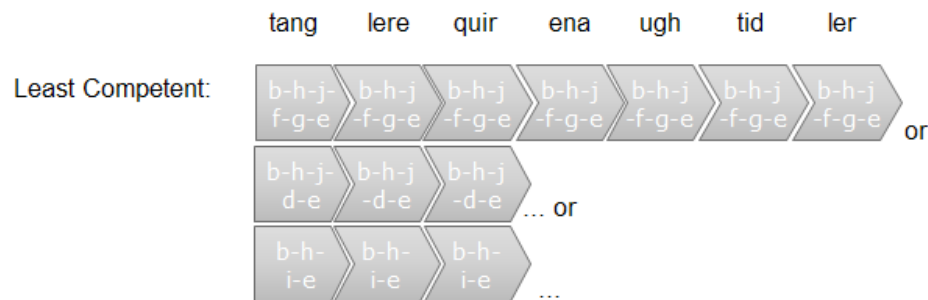


Figure 4.7 Example of the assumed chunking at syllable level and the processes that might take place during the sentence-copying

These examples provide some idea of the different chunking possibilities that might occur at different levels of processing. The next section shows a detailed simulation example using the sentences used in Experiment 1. The application: Sentence Simulation Example

This section demonstrates how the process of simulating the sentences takes place, using three different types of sentences:

Example 1: Jumbled Proverb with high frequency words

‘are than bigger eyes your belly ’

Example 2: Sentence with low frequency words and omission of spaces between words

‘tanglerequirenaughtidleredictobtusenuptialepisode’

Example 3: Sentence with Jumbled Letters

‘tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef’

Figure 4.8 Example of the assumed chunking at syllable level and the processes that might take place during the sentence-copying

The above sentences have been manipulated in terms of word presentation and meaning. The difficulty participants will have in processing these sentences, and the resulting pause lengths, will vary depending on the manipulation used. The main objective of sentence manipulation is to increase the differentiation between the varying levels of nature in language processing.

It is predicted that:

- Example 1, a jumbled proverb, will disturb the overall comprehension process. The sentence however, can be processed at word level chunking by both the MC and the LC, as it contains short, high frequency, familiar words. If the proverb is familiar to the MC, it can be processed in bigger chunk sizes, at group level. If the proverb is not familiar to the participant, it will be processed like a word-list.
- Example 2, which contains low frequency words with the omission of spaces between them, makes it difficult to recognise words. Participants will process this kind of sentence at syllable level, where long pauses can be observed more frequently within a word.
- In example 3, a sentence with jumbled letters in the words, word recognition will be more difficult, and hence long pauses at syllable level will be more likely. The first and last letters of the words are preserved, and only the middle letters are jumbled. The purpose of manipulating the letters in such a way is to make word recognition difficult. A competent person might find this difficult to copy because of the letter transposition and will therefore produce long pauses at both word level and syllable level. Moreover, the irregularly spelt words will make copying difficult for the MC because of the conflict between the instinctive need to correct the spelling and the instruction to copy exactly what is being shown.

The next step is to simulate these sentences with the theoretical model (MoC), guided by the information presented earlier in this chapter.

Figure 4.9 below explains the features that will be used to predict the chunk occurrences, which applies to Figures 4.10. The numbering given in Figure 4.9 describes the functions as listed below the figure.

Below are the guides (represented by the numbers) to understand the functions of each cell in Figure 4.9:

- With the above guidance, Figure 4.10 below demonstrates the chunking processes and the PLP values that are assigned based on the word characteristics. The proposed model provides some direction to make sure that the design of the sentences can differentiate between levels of competence. Example 2, as can be seen, differs from the rest because of the nature of the sentence manipulation that omits the spaces between words, hence there are no black vertical borders to separate the words. With Example 1, there is a clear chunk between words, however,

with Example 3, because of the jumbling of letters condition, has shown some possibility of syllable chunking. Overall, the three examples provide three different sentence characteristics that might be used as a means to measure the different cognitive processes in language processing.

To illustrate this further, the PLP values plotted in Figure 4.10 were mapped onto a line graph, Figure 4.11, to observe the PLP values distributions across the successive letters. The data show that overall; LC produces longer PLP values compared to MC, who produces shorter PLP values. With Example 1, MC is able to perform larger chunking size compared to LC, but the case is different with Example 2 and 3, where the chunking sizes of MC and LC is similar. This may be because of the difficulty to copy the manipulated sentences. Overall, the median of PLPs obtained as in Table 4.5 shows an increase from Example 1-3, which might indicate the difficulty of processing the sentence in Example 2 and 3.

	Example 1	Example 2	Example 3
Most Competent	4	4	5
Least Competent	5	6	6

Table 4.5 Median of PLPs for the three examples

Example 1

Words Jumbled	a	r	e	t	h	a	n	b	i	g	g	e	r	e	y	e	s	y	o	u	r	b	e	l	l	y
MC Chunking	4													4												
Word Group / Word / Syllable	4			1				1						4				1				1				
Letter		1	1		1	1	1		1	1	1	1	1		1	1	1		1	1	1		1	1	1	1
Words Jumbled	a	r	e	t	h	a	n	b	i	g	g	e	r	e	y	e	s	y	o	u	r	b	e	l	l	y
LC Chunking	4							5						5				5				5				
Word Group / Word / Syllable	4			1				5						5				5				5				
Letter		1	1		1	1	1		1	1	1	1	1		1	1	1		1	1	1		1	1	1	1

Example 2

Words Jumbled	t	a	n	g	l	e	r	e	q	u	i	r	e	n	a	u	g	h	t	i	d	l	e	r	e	d	i	c	t	o	b	t	u	s	e	n	u	p	t	i	a	l	e	p	i	s	o	d	e			
MC Chunking	4				4				4					4					4				4					4					4				4				4											
Word Group / Word / Syllable	4				4				4					4					4				4					4					4				4				4				4							
Letter		1	1	1		1	1	1		1	1	1	1		1	1	1	1		1	1	1		1	1	1	1		1	1	1	1		1	1	1		1	1	1		1	1	1		1	1	1		1	1	1

Words Jumbled	t	a	n	g	l	e	r	e	q	u	i	r	e	n	a	u	g	h	t	i	d	l	e	r	e	d	i	c	t	o	b	t	u	s	e	n	u	p	t	i	a	l	e	p	i	s	o	d	e		
LC Chunking	8				6				6					6					8				6				8				6				6				8				6								
Word Group / Word / Syllable	8				6				6					6					8				6				8				6				6				8				6								
Letter		1	1	1		1	1	1		1	1	1	1		1	1	1	1		1	1	1		1	1	1	1		1	1	1		1	1	1		1	1	1		1	1	1		1	1	1		1	1	1

Example 3

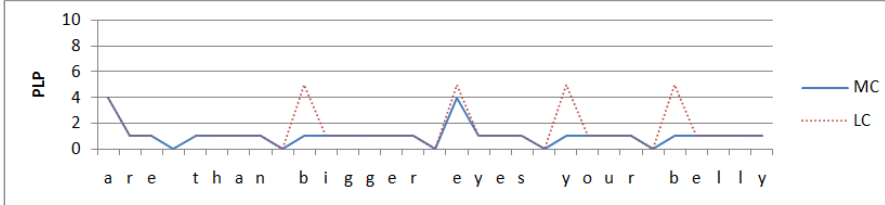
Words Jumbled	t	i	h	s	i	s	b	c	u	s	e	a	e	t	h	e	h	u	a	m	n	m	n	i	d	d	e	o	s	n	o	t	r	a	e	d	e	r	v	e	y	i	t	e	t	e	r	b	y	i	s	t	l	e	f
MC Chunking	5				4		5			6				5			5					5				5			4			5			5				5				5			4			5						
Word Group / Word / Syllable	4				4		5			6				5			5					5				4			4			4			5				5				5			4			5						
Letter		1	1	1		1			1	1			1	1	1		1	1	1	1			1	1	1		1	1	1		1	1		1	1	1		1	1	1		1	1	1	1		1		1	1	1	1		1	

Words Jumbled	t	i	h	s	i	s	b	c	u	s	e	a	e	t	h	e	h	u	a	m	n	m	n	i	d	d	e	o	s	n	o	t	r	a	e	d	e	r	v	e	y	i	t	e	t	e	r	b	y	i	s	t	l	e	f		
LC Chunking	6				5		8			6				5			8			6			7		6		6			5			6			8			6			7			6			5			8			6			
Word Group / Word / Syllable	6				5		8			6				5			8			6			7		6		6			5			6			8			6			7			6			5			8			6			
Letter		1	1	1		1			1	1			1	1	1		1	1		1	1	1		1	1	1		1	1		1	1		1	1		1	1	1		1	1	1		1	1	1		1		1	1	1		1		

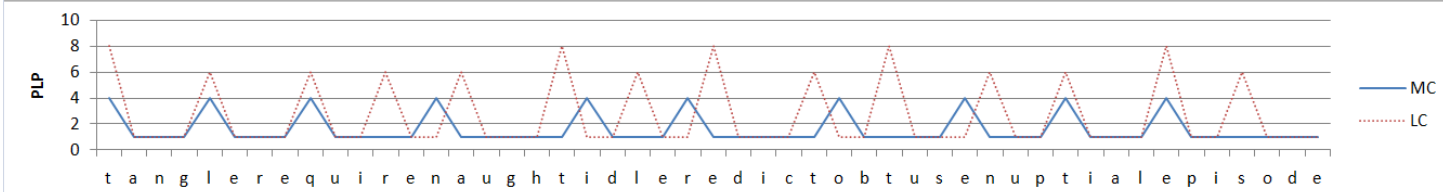
Figure 4.10 Examples of PLP Plotting

[Key: MC: Most Competent; LC: Least Competent; L2: Group of Words (Phrase)/Word level chunking; L1: Syllable/Letter level chunking]. The values in the boxes are the PLPs.

Example 1



Example 2



Example 3

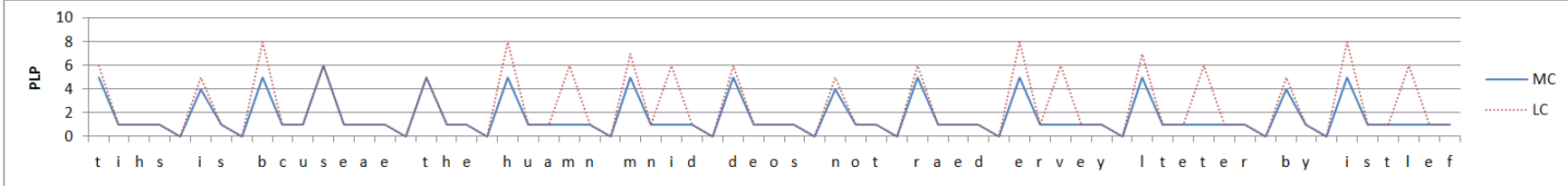


Figure 4.11 Examples of real data and PLP data plotted on a line graph

The way the data are structured in Figure 4.11 enhances the understanding of the copying processes, in addition to drawing attention to how these differ between participants of different competence levels. Designing sentence stimuli is not a straightforward task. Running a sentence through the MoC might therefore help to determine whether a given sentence is suitable for assessing different competence levels. Another way of analysing this prediction is by observing the chunking patterns by the number of peak value points. The bigger the differentiation between the MC and the LC, the more suitable the sentence is likely to be.

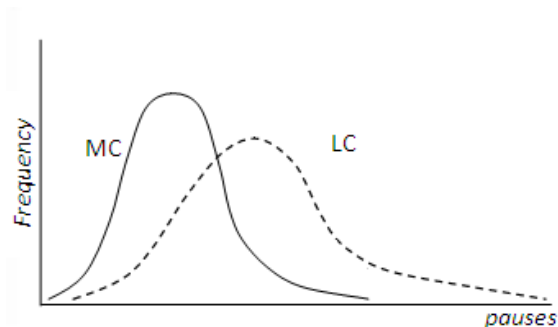


Figure 4.12 An illustration of the pause distribution differences between MC and LC that the thesis aims to uncover

Figure 4.12 illustrates an example of the difference of pause distributions between MC and LC. These data help this research in ensuring suitable sentence stimuli are designed, that will effectively measure the cognitive processes of language processing. The manipulations in the sentences influence the chunking size and the pause lengths produced. For instance, a difficult sentence such as Example 2 and 3, might cause less competent individual to produce long pauses, so the LC graph is stretched, compared to the MC. There might also be the possibility of a swap between MC and LC, where the sentence manipulation affects MC to produce longer pauses than LC. This may be the case of sentences that requires the involvement of comprehension. The different levels of familiarity of the stimuli encountered would definitely affect pause length. The wide span of the bell curve shows that pause lengths vary at different levels – letter, syllable, word or word group. There might be the possibility that the long pauses are occurring at syllable level chunking and short pauses are occurring at word level chunking. Therefore, the kind of statistical measure chosen to measure these pauses is an important aspect for exploration. In brief, Figure 4.12 is an example of a graph that assists in illustrating how a sentence is effective in differentiating the levels of competence, which indirectly provides the answer to the suitable sentence and measure.



## 4.5 Conclusion

This chapter introduced the theoretical Model of Copying (MoC) used for understanding the different processes involved in the activity of copying, for designing the stimuli, and for interpreting the data. It described how the model was developed and explained the features of the model, as a guide to using it.

The model presents seven different routes that might be used by individuals in the process of copying sentences, and is predicted to be able to explain the different pause lengths produced. These routes will differ according to the nature of language processing of the participants and the characteristics of the words used in the sentence stimuli. As a way of turning the assumptions into specific predictions, the 'processing load prediction' (PLP) was introduced. The PLP also helps with the simulation of sentences, by predicting chunk sizes. These predictions are useful for the design of sentence stimuli.

The motivation for producing the MoC is due to the need to understand the processes involved in copying; hence the guidelines for using the model as presented in this chapter. However, the scope did not cover validating the model; a considerable amount of extra time would be needed for this. A number of limitations were discovered within the development of this model: the effect of parallel processing, *strategy effects*, as defined by Coltheart (1978), and the possibility of practice effect. These are considered below.

Parallel processing is characterised by the possibility that the reading process can be simultaneous to the handwriting production, especially in competent participants. The greater familiarity with the stimuli results in extra processing resources in the working memory, which allow them to have multiple sub-processes occurring at the same time. For example, a competent person might have lower processing costs with regards to handwriting production because of their greater familiarity with writing the scripts and with regards to words and sentence-related processes because of their greater familiarity with the language. Thus, it becomes possible that while copying, the competent person can also be reading and processing the chunking of the following words. This occurrences of the parallel processing results in flatter pauses in the process of copying, which means that for some processes (in the MoC), they occur simultaneously, which may appear like a loop until the copying is completed.

The *strategy effect* as claimed by Coltheart (1978) is said to occur when participants are faced with nonword as opposed to regular words as they tend to slow down in reading especially when the nonword appears first before the regular word. Although Coltheart has found this in his study of *reading aloud* which has resulted in the DRC Model, it was predicted that similar occurrences might occur with regards to word recognition. In the case of this thesis work, it was assumed that participants would give up on attaining comprehension therefore look for a simpler strategy to complete the copying. Without realising, they will eventually turn away from the lexical route. This would lead them to the non-lexical route for the purpose of completing the task. Such changes of route would increase the pause lengths for both MC and LC.

As a conclusion, it is hoped that the model will benefit this research work in terms of understanding the underlying cognitive processes and in search of the suitable sentence stimuli and measure of the cognitive processes in language processing.

The next chapter will continue the investigation, following the findings of the Pilot Study, to establish the most effective sentence type, using IC as the method of copying, and the median of pause levels as the measure of cognitive processes. It will focus on L2 word pauses, which are more meaningful than L1 pauses. The present methodology will be improved, with the introduction of a more reliable independent English test and an increased number of participants from a more focused group.

## CHAPTER 5      **Experiment 1: Investigating Suitable Sentence Stimuli to Assess the Cognitive Processes of Language Processing**

### 5.1 Introduction

This chapter aims to (1) investigate the suitable sentence types for differentiating the varying competence levels of language users, and (2) explore the role of pauses in order to get some information about the cognitive processes. There are many questions raised by the pilot study which still need to be addressed. As such, this chapter will (1) continue examining IC, to confirm it as an effective method of copying and (2) test Median of pauses (i.e. L0, L1 and L2) as an effective statistical measure, focusing on L2 word pauses and (3) continue examining the sentences used in the pilot study.

Firstly, in the investigation of effective sentence stimuli, the same sentences as those applied in the pilot are used in this study, specifically those used with the IC condition. A sentence stimulus is judged as effective if it is able to differentiate levels of competence, which is shown through the greater correlation values between the pauses captured and an independent form of language test. An effective stimulus will be able to demonstrate that a competent participant with good knowledge of English will be able to copy the sentence fluently; producing short L2 (word) level pauses. With the same stimulus, a less competent English language user is expected to find difficulty in copying at word level, because of the unfamiliarity of the words and sentence structure, therefore producing long L2 (word) level pauses.

The second aim of this chapter is to explore the information that pauses are able to reveal regarding the nature of language processing, where pauses consist of processes occurring at L2 (word), L1 (letter) and L0 (stroke) level. This study explores how these different levels of pauses could help in explaining the patterns of chunking and processing go on during the course of copying. The pause lengths signify the amount of processing required to perform the copying of the sentence. Long pauses demonstrate a greater processing demand of the working memory, while short pauses demonstrate a lesser processing demands. Generalisations about an individual's competence cannot be made on the basis of pause length, because it depends on the type of sentence used, the manipulation employed and language background of the individual.

However, examining pauses at different levels is able to provide information on the kind of processing that might have taken place at a particular point in time.

The study aims to establish which pause level is effective for measuring the cognitive processes of language processing. Although theoretically, L2 (word) level pauses are more meaningful than L1 (letter) level, the latter showed better correlation values in the pilot. This is therefore further examined in this chapter, to confirm which pause level is more meaningful.

Capturing pauses in baseline copying serves a different purpose than capturing pauses in sentence-copying. The baseline aims to capture pauses that represent the processing of basic level writing, which should be a greatly familiar task for all participants. Establishing the baseline pauses enables the researcher to correct any pause data unrelated to the nature of language processing, such as data related to the individual effects of speed writing. With the assumption that letter writing within the sentence-copying is of great familiarity to all participants, it ought to be possible to use the L1 (letter) pauses within the sentence-copying for normalising. This chapter aims to explore the role of these pauses further.

In order to confirm IC as an effective method of copying and median of pauses (i.e. L0, L1 and L2) as an effective statistical measure, this experiment will investigate the same variables in order to confirm the findings of the pilot study. This investigation is considered important as the pilot only included five participants, with various backgrounds. This experiment will carry out further tests on a better group of participants to (1) confirm the suitability of the findings and (2) identify the appropriate sentence stimuli for measuring the cognitive processes of language processing.

It is vital, however, to take into account that ESL participants can be subject to a variety of First Language (FL) effects, such as the orthographies used in their FL, e.g. Chinese or Arabic characters. Since it was considered crucial in this study to restrict the category of participants in a more uniform manner, it was decided that a population be whose first languages all use the same script as the English language. 20 Malaysian participants were therefore recruited. Changing the orthographic features in any writing task bears a processing cost, which needs to be minimised. As this study is interested only in the language-related processes during the course of copying, it

would be beneficial to exclude as many unrelated processing costs as possible. Hence, the study recruited only adults, because their handwriting production should already be fluent.

The pilot study has highlighted the need for a reliable independent measure to obtain information on the most recent language performance of the participants, to validate whether the sentence-copying is really measuring the cognitive processes. This experiment therefore applies the most frequently used statistic for validating tests in Language Testing (LT), the correlation coefficient, because of its ability to show whether both the test and sentence-copying are measuring similar aspects of language knowledge (i.e. components). Instead of ranking participants based on their biographic information and interviews, as in the pilot, scores obtained from an independent English test were used, as they were considered to be a more reliable approach – the interpretation of the former method can be imprecise.

Summing up, it was predicted that participants with good English scores, being more familiar with the language, will be able to produce shorter L2 (word) level pauses during the copying of English sentences. By contrast, participants with lower English scores will be less familiar with the language, hence producing longer L2 (word) level pauses.

## 5.2 Method

### 5.2.1 Participants

Twenty Malaysian participants with an age range between 20 and 30 took part in this study. None had language impairments or handwriting problems. The participants all had English as their second language (ESL) and had completed their secondary education qualification (Sijil Penilaian Malaysia). English language is a compulsory subject in Malaysian schools and is introduced to all schoolchildren from kindergarten. Further, English is commonplace on television, in newspapers, magazines and books, and is used extensively within organisations. The multi-ethnic nature of Malaysian society encourages English usage, which explains why it has been chosen as the second language in the country.

### 5.2.2 Apparatus

Similar tools were applied to those used in the pilot study, explained in Section 4.2.2. In addition to these, however, the research in this chapter took advantage of a program called “PLET – Pause

*Length Extraction Tool*” (van Genuchten, 2009). This was used to code and conduct the preliminary analyses of the data, which includes identifying pause levels (i.e. L0, L1 and L2).

### 5.2.3 Procedure

Overall, the same procedure was followed as in the pilot study. The participants were asked to copy the sentence stimuli as quickly and accurately as possible. The stimuli were shown one at a time, and the process was repeated until all the stimuli had been copied. The participants were not allowed to stop during the course of copying for any reason. Each stimulus was visible at all times whilst being copied. The participants, however, were not allowed to read them in advance; they were only allowed to read the stimuli and begin copying when the researcher said ‘start’. The participants were trained to begin each sentence with a ‘hash’ (#) in order to ensure that the process of writing was well underway before the first stimulus letter was copied. Each participant wrote 6 baselines and copied 19 sentence stimuli in the order given (see Table 3.1). Upon completing the writing tasks, they were asked to take a general test of language proficiency (as an independent measure of cognitive processes) which is explained in Section 5.2.5. In total, the 20 participants copied 25 stimuli each, producing 500 data records.

### 5.2.4 The Stimuli: Sentences & Baselines

The same stimuli used for IC in the pilot study were replicated in this experiment (see Figure 4.2 and Table 3.1). . In this experiment, a total of 25 stimuli had to be copied; 19 of those were manipulated English sentences and 6 were baselines. The baselines consisted of writing letters of the alphabet, combinations of 2, 3, 4 and 5 random letters and the participant's own name (Figure 4.2). The manipulated sentences included sentences on technical topics, sentences with jumbled words, sentences containing words with jumbled letters, and letters of the alphabet in correct and reverse orders (Appendix 2). Detailed explanations of each manipulation are given in Section 4.2. All participants were required to copy the baselines and sentence stimuli in the same order (from 1 to 25).

### 5.2.5 The Independent Measure of English Language

To assess the participants’ general language competence, an independent measure of English language abilities was introduced, in the form of a free online English grammar test published by Oxford University Press (2009), the Oxford Diagnostic Test (ODT). Although this online test is not a

standardised psychometric test, it is considered a suitable means to validate the present experiment because it uses modes of response that are unrelated to the pauses, in particular multiple choice and typing into blank fields. The ODT tests a wide variety of language knowledge skills such as tenses, modals, passive constructions, negatives, articles, determiners, prepositions and conditionals. Moreover, it has three levels of difficulty (basic, intermediate and advanced) each comprising 100 questions, which are equally weighted. The participants' results from the ODT were scored on a percentage scale and then correlated to the pauses.

### 5.2.6 The Overall Structure of this Experiment

Figure 5.1 shows the overall structure of this experiment.

Four aspects presented at level C group a number of variables (level D) that are important to this study. Some of these variables were established in the pilot, and some are developed in this chapter. Out of these variables, those used in previous experiments are (1) The method of copying (Immediate Copying (IC) has now been chosen), (2) the baseline stimuli, (3) the sentence stimuli used for IC and (4) the medians of pauses (i.e. L2s and L1s). The additional variables introduced in this chapter are (1) the language proficiency test (i.e. ODT) and (2) the normalisation.

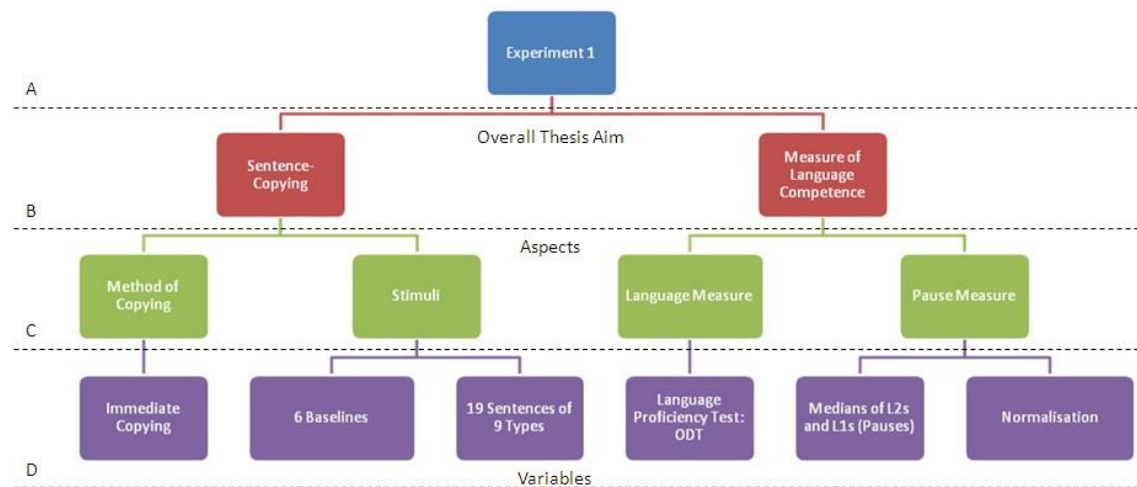


Figure 5.1 The overall structure of this experiment

### 5.2.7 Analysis

The raw data captured comprise 6 baseline writings and 19 sentence stimuli, making a total of 25 stimuli for each of the 20 participants. The data from copying these stimuli were then extracted using TRACE in order to obtain the pauses between each stroke. Instead of mapping each pause value to the letters of the sentences and then determining the pause levels manually, PLET (van Genuchten, 2009) was used to locate the pauses and perform the calculation. After obtaining all the data, they were then correlated with the ODT scores in order to achieve the aim of this chapter. The levels of analysis are explained in the results section. The analyses were conducted in a top-down fashion, from the greatest aggregation down to specific cases:

1. Effects of the different ODT test levels.
2. Effects of the ODT and the overall pause values (L2, L1 and L0).
3. Effects of the ODT and the aggregated pause levels.
4. Effects of using L1 from baseline writing or L1 from sentence stimuli, as a means of normalising.

For the purpose of finding the effective sentence stimuli, the sentences were then refined based on the normalised data.

## 5.3 Results

### 5.3.1 Online Oxford Diagnostic English Test Level

The ODT consists of three levels of difficulty: Basic, Intermediate and Advanced. The mean score for the Basic ODT level was 76.5% (SD=9.84%), for the Intermediate level was 75.25% (11.21%) and for the Advanced level was 58.9% (10.76%). The mean scores for Basic and Intermediate are substantially higher than Advanced, indicating that Advanced is more difficult.

Three paired t-tests were performed in order to examine the overall consistency of the measures provided by the three levels of the test. These showed a significant difference between the scores in the Basic and Advanced levels:  $t(19)=11.48$ ,  $p<.0001$ , and between the Intermediate and Advanced levels:  $t(19)=9.92$ ,  $p<.0001$ . There was no difference, however, in the scores between the Basic and Intermediate levels:  $t(19)=0.96$ ,  $p=0.35$ .



Correlation between the 3 tests:	
Basic & Intermediate	0.854*
Basic & Advanced	0.782*
Intermediate & Advanced	0.776*

Table 5.1 The correlation values between the three tests of ODT (\*p<.001)

A Pearson correlation was also conducted, which is presented in Table 5.1. The correlations between the tests are significantly strong, with the overall value of above 0.7 (p<.001). This confirms that the three tests are measuring similar components of language.

Overall, the means of scores denote that the three tests agree with the original rank, from Basic being the easiest, to Advanced being the most difficult. The t-test indicates no difference between the Basic and Intermediate (also supported by the strong correlation gained), which might signify that using Basic and Intermediate is redundant, because doing either one of these two tests might provide similar results. However, the Advanced test is able to differentiate the levels of competence better than either the Basic or Intermediate tests. To conclude, the strong correlations also confirm that the application of ODT is a suitably independent measure of language competence.

### 5.3.2 Language Performance versus Overall Pause Value

Table 5.2 presents the results obtained from performing the correlation between the ODT and the overall pause value (L2, L1 and L0). It is apparent from this table that the median provides weak negative correlations, but the mean provides significant negative correlations. The weak negative correlations of the median do not agree with the pilot findings, suggesting the need for further attempts to investigate each pause level separately.

		Measure	
		Median	Mean
ODT Tests	Basic	-0.256	-0.488*
	Intermediate	-0.334	-0.568**
	Advanced	-0.117	-0.468*
	Mean Score	-0.174	-0.546**

Table 5.2 Correlations between the Overall Pauses of L2, L1 and L0 against the ODT tests (\*p<.05, \*\*p<.01 and \*\*\*p<.001)

### 5.3.3 Language Performance Versus Pause Levels

In order to explore the kinds of information that can be revealed from the pauses, the analysis was performed individually on each of the median pause values for L2 (word), L1 (letter) and L0 (stroke) levels for each participant on each task. The means of those median pauses were calculated for each participant across the tasks: L2=485ms (SD=163.2s), L1=290ms (87.1s) and L0=90ms (14.7s). T-tests were also applied to each pair of pause levels and were all found to be statistically significant: L2 and L1,  $t(19)=6.03$ ,  $p<.0001$ ; L2 and L0,  $t(19)=11.24$ ,  $p<.0001$ ; L1 and L0,  $t(19)=11.5$ ,  $p<.0001$ . The pause length clearly increases as the chunk level increases from L0 to L2. This is consistent with previous studies (c.f. Chapter 2), which have argued that pause values reflect the different levels of the hierarchy of chunks.

These pauses at the L2, L1 and L0 levels were then correlated with the ODT tests (see Table 5.3) to examine the pause levels that measure the cognitive processes.

ODT Tests	Pause Levels		
	L2 (Word)	L1 (Letter)	L0 (Stroke)
Basic	-0.585*	-0.303	-0.204
Intermediate	-0.681**	-0.371	-0.139
Advanced	-0.649**	-0.161	-0.313
<i>Mean of correlation</i>	<i>-0.687**</i>	<i>-0.210</i>	<i>-0.243</i>

Table 5.3 Correlations between ODT scores and pause durations across writing levels (\* $p<.01$  and \*\* $p<.001$ )

The correlation values for L2 level display significant negative correlations, but not at L1 and L0 levels. The significant negative correlations observed at the word chunk level suggest that participants may be processing the sentence-copying more meaningfully at L2 (word) level than at L1 (letter) and L0 (stroke) level. The negative correlation shows that a competent participant, who scores well on the ODT, produces shorter L2, L1 and L0 pauses, whereas a less competent participant, with poor ODT scores, produces longer L2, L1 and L0 pauses. The fact that the correlation for L1 and L0 is weak may be explained by the fact that these levels do not appear to be meaningful in relation to language and thus do not signify the cognitive processes of language processing. It is important to note that the correlation values do not vary consistently with the ODT tests.

### 5.3.4 Normalisation: Correcting Individual Writing Speed

Here, an attempt is made to explore ways to compensate or correct the speed of writing, by first presenting the baseline results, followed by the technique of normalising.

#### 5.3.4.1 The Baseline

The purpose of the baseline tests is (1) to capture the basic writing skills of highly practised letters and most frequently written words (i.e. name writing) in order to investigate if copying or writing these has any relation to the cognitive processes of language processing, and (2) to correct the raw data for any effects of speed writing. A summary of the findings is presented in Table 5.4. Based on the t-test findings presented in Section 5.3.1, this focuses specifically on the Advanced level of the ODT test, considered to be more reliable than the Basic and Intermediate levels.

In the investigation of any relation between the L2 and L1 baseline pauses and language competence, all the stimuli show weak correlations except for the ‘upper case name writing’, which showed a strong correlation with the L2 baseline. The weak correlations are a good indication that writing the letters of alphabet reveals nothing about the writer’s nature of language processing. However, the fact that the correlation is higher ( $r(18)=-0.605$ ,  $p<.01$ ) for ‘upper case name writing’ is interesting, and might suggest that there is some relation to language processing. This will be discussed further in Section 5.4.

<b>Correlations</b>	Alphabet Letters in Upper Case	Alphabet Letters in Lower Case	Jumbled letters with varying Upper and Lower cases	Combination of letters to form word-like Sentences	Upper case Name Writing	Lower case Name Writing
L1 Baseline versus ODT	-0.329	-0.316	0.009	-0.288	-0.158	-0.184
L2 Baseline versus ODT				-0.106	-0.605**	-0.325
L1 Baseline versus L1 Sentences	0.689***	0.848***	0.533*	0.820***	0.904***	0.944***
L2 Baseline versus L2 Sentences				0.542**	0.602**	0.828***

Table 5.4 Summary of all correlations between L2 and L1 pauses against ODT and Sentence Stimuli (\* $p<.05$ ; \*\* $p<.01$ ; \*\*\* $p<.001$ )

In the investigation to see the effects of L2 and L1 baseline pauses on the sentence-copying, strong positive correlations were found for all the baseline stimuli at both L2 and L1 levels. This suggests that copying highly practised letters and name writing in the baseline uses similar processing to that used in copying the letters and words in the sentence-copying. This also suggests that using L1 pauses from the sentence-copying may be sufficient for normalising (this will be explained further in the next section). Figures 5.2 and 5.3 further illustrate the correlations between each baseline and each sentence stimuli.

In Figure 5.2, 'lower case name writing' shows a strong correlation across all sentences for L1 (letter) level. The lowest correlations are for 'jumbled letters with varying upper and lower cases', which is not surprising because of the manipulation of letter position and letter presentation. This suggests that the manipulation employed interrupts the fluency of copying the correct order of letter of alphabets.

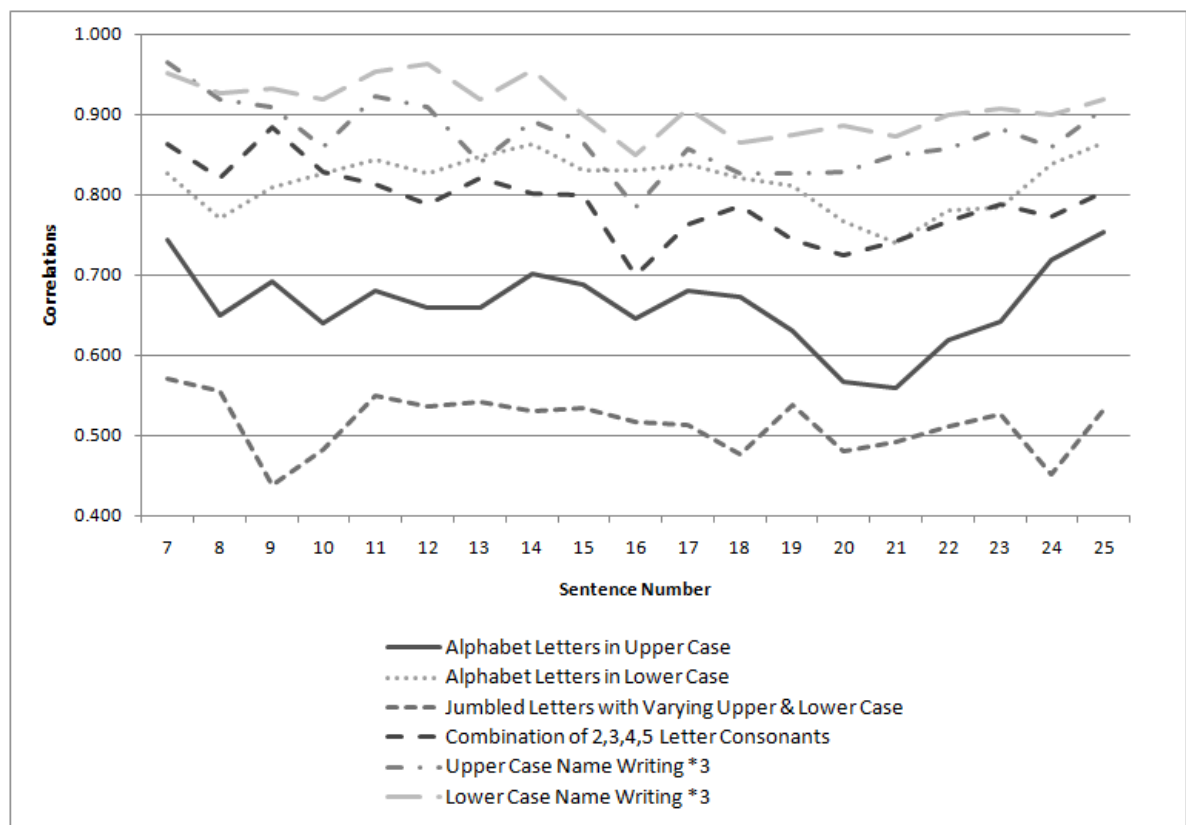


Figure 5.2 Baselines correlated against the sentence stimuli at Letter (L1) Level

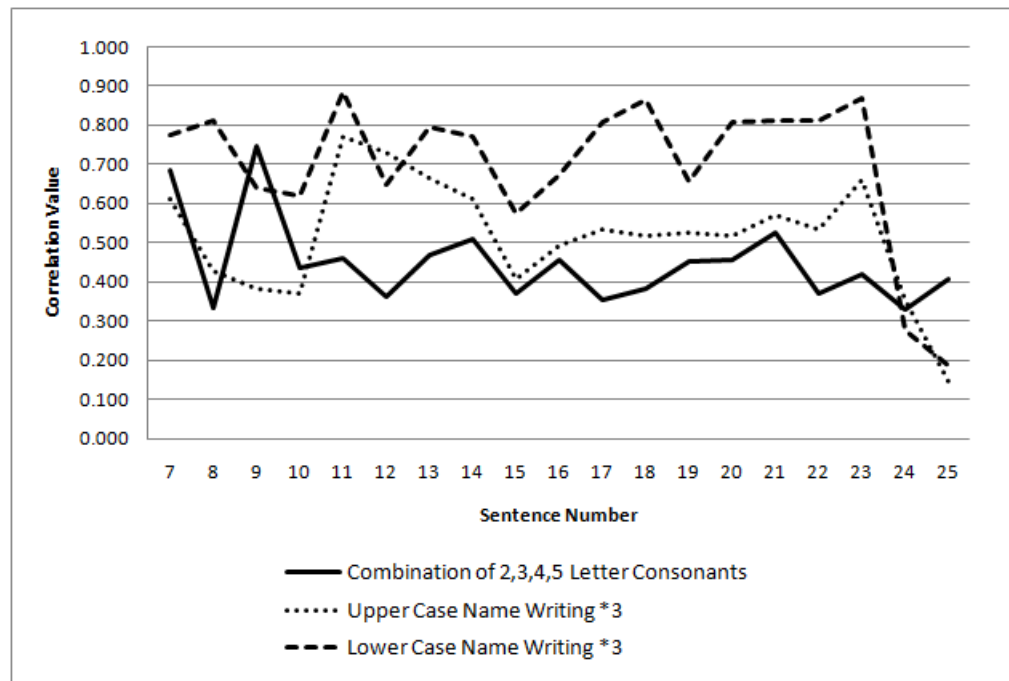


Figure 5.3 Baselines correlated against the sentence stimuli at Word (L2) Level

The graph in Figure 5.3 illustrates the correlations between the L2 baselines and the L2 sentences. These correlations vary across the sentences. Writing names in lower case shows a stronger correlation than writing names in upper case. However, the most striking result to emerge from the data regards sentence numbers 24 and 25, in which the reaction to the correlation differs from the rest of the sentences. When these sentences are compared with the graph in Figure 5.2, L1 (letter) level pauses show a stronger correlation, than L2 pauses. The lower correlation found at L2 (word) level pauses for the two sentence might reflect the manipulation used in the sentence stimuli.

#### 5.3.4.2 The Normalisation Technique

Two approaches were used in the attempt to compensate for the discrepancy in individual writing speeds. Both involved the use of L1 (letter) pauses during the copying of sentence stimuli. As explained in the previous section, this approach to normalisation only focuses on the L1 pauses from the sentence-copying, and not the L1 pauses from the baseline. The two approaches are the

*pause quotient* and the *pause difference*. The *pause quotient* is calculated by dividing the mean of L2 (word) median pauses for all the sentences for each participant by the mean of L1 (letter) median pauses for all the sentences for each participant (L2/L1). *Pause difference* subtracts the mean of L1 (letter) median pauses for all the sentences for each participant from the mean of L2 (word) median pauses of all the sentences for each participant (L2-L1). The mean value across all participants for *pause quotient* is 1.75 and for *pause difference* is 195s.

The correlations between these measures and the ODT scores are shown in Table 5.5. The ODT Basic and Intermediate data are also presented to in order to support the choice of using only the ODT Advanced data to validate the method. When comparing the values in the table, the absolute magnitude of the correlation increases with the Advanced ODT, but this is not the case for the Basic and Intermediate ODT scores. The order of magnitude of the correlations (after the normalisation with both approaches) now follows the order of the ODT tests. This suggests that there may be some merit to using L1 pause values from the sentence-copying to account for individual writing speeds when measuring correlations.

		L2 Median	Normalisation	
			Pause quotient (L2/L1)	Pause difference (L2-L1)
ODT Tests	Basic	-0.585**	-0.459*	-0.527**
	Intermediate	0.681***	-0.525**	-0.587**
	Advanced	-0.649**	-0.704***	-0.690***
	Mean Score	0.687***	-0.606**	-0.647**

Table 5.5 Correlations between ODT scores and the Pause Quotient and Pause Difference (\*p<.05; \*\*p<.01; \*\*\*p<.001)

### 5.3.5 Refining the Tasks

Normalisation has helped in correcting the effect of less competent participants who simply wrote faster than other participants (i.e. outliers). The results in Table 5.6 below show the correlation of each stimulus (median of L2 pauses) with the Advanced ODT scores after being normalised, showing that the normalisation has either increased or decreased the original correlation values.

In trying to identify the most effective sentence stimulus, the sentences are arranged in Table 5.6 below, in descending order from the highest significant correlation value to the lowest. This shows that the correlation value of each individual sentence is not consistent between sentences of the same group (represented by the letters of the alphabet). The order in which the sentences were presented to the participants does not seem to have affected the copying process either: the order of the sentence numbering bears no relation to the ranking of the correlations values. This presents the opportunity to increase the overall accuracy of the measures of cognitive processes in language processing, by selecting the tasks with the highest correlations. The sentences are therefore grouped based on the strong correlations achieved.

Group	No	Sentence Description	Method A: (L2Median)	Normalise	
				Pause Quotient (L2/L1)	Pause Difference (L2-L1)
<b>E</b>	17	Garden path sentence	<b>-0.702***</b>	-0.736***	-0.732***
<b>G</b>	21	Two proverbs combined by alternately selecting words	<b>-0.691***</b>	-0.655**	-0.656**
<b>E</b>	15	Garden path sentence	<b>-0.684***</b>	-0.674**	-0.649**
<b>D</b>	14	Words with jumbled letters, except for the end-letters	<b>-0.666***</b>	-0.737***	-0.710***
<b>B</b>	10	Technical sentences with specialist terms from specific fields and no spaces in between words	<b>-0.664***</b>	-0.647**	-0.658**
<b>D</b>	13	Words with jumbled letters, except for the end-letters	<b>-0.660***</b>	-0.749***	-0.748***
<b>A</b>	8	Technical sentences with specialist terms from specific fields	-0.619**	-0.636**	-0.649**
<b>H</b>	22	Three proverbs combined by alternately selecting words	-0.618**	-0.631**	-0.612**
<b>F</b>	19	Common words, with no spaces in between. The last and first letters of successive words are high frequency bigrams	-0.618**	-0.602**	-0.611**
<b>E</b>	16	Garden path sentence	-0.584**	-0.613**	-0.565**
<b>F</b>	18	Common words, with no spaces in between. The last and first letters of successive words are high frequency bigrams	-0.572**	-0.510*	-0.552**
<b>G</b>	20	Two proverbs combined by alternately selecting words	-0.567**	-0.615**	-0.589**
<b>H</b>	23	Three proverbs combined by alternately selecting words	-0.556**	-0.637**	-0.617**
<b>C</b>	11	A proverb with jumbled words	-0.550**	-0.639**	-0.686***
<b>A</b>	7	Technical sentences with specialist terms from specific fields	-0.416*	-0.448*	-0.466*
<b>C</b>	12	A proverb with jumbled words	-0.398*	-0.427*	-0.432*
<b>B</b>	9	Technical sentences with specialist terms from specific fields and no spaces in between words	-0.345	-0.187	-0.221
<b>I</b>	24	Low frequency words, with no spaces in between. The last and first letters of successive words are high frequency bigrams.	-0.178	-0.148	-0.124
<b>I</b>	25	Low frequency words, with no spaces in between. The last and first letters of successive words are high frequency bigrams.	-0.100	-0.021	-0.048

**Table 5.6 Correlations for each sentence stimuli with Medians of L2s and Normalised data in search of the effective sentence (\*p<.05, \*\*p<.01 and \*\*\*p<.001)**

Tasks number 8, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 and 23 (i.e. the best 14) were found to have statistically significant correlations, of more than  $r(18) = -0.55$  ( $p < .01$ ). Of those, tasks

number 10, 13, 14, 15, 17 and 21 (i.e. the best 6) were found to be statistically significant, with more than  $r(18)=-0.66$  ( $p<.001$ ). Rather than aggregating the L2 pause values over the entire set of sentence stimuli data, these two subsets (best 6 and best 14) can now be used in search of the effective sentence type, with particular focus on the subset with the strongest correlation (best 6).

Correlations with ODT	L2 Medians	Pause Quotient (L2/L1)	Pause Difference (L2-L1)
All 19	-0.649*	-0.709**	-0.690**
Best 14	-0.699**	-0.746**	-0.730**
Best 6	-0.712**	-0.747**	-0.726**

**Table 5.7 Correlations between ODT Advanced and the refined pauses, grouped by the correlation value, before and after normalisation (\* $p<.01$  and \*\* $p<.001$ )**

Table 5.7 shows the effect of eliminating the weaker stimuli on the normalisation calculation performed on the pause data. These sentences are grouped in order of the most significant correlation values, as described previously. ‘All 19’ includes all the sentences. ‘Best 14’ comprises the sentences with the 14 highest correlations above  $-0.55$  ( $p<.01$ ) and ‘best 6’ refers to the top six sentences, i.e. sentence numbers 10, 13, 14, 15, 17 and 21, which have correlation values above  $-0.66$  ( $p<.001$ ). The correlation increases gradually from ‘All 19’ to the ‘Best 6’, (except for the slight difference between *best 14* and *best 6* for *pause difference*) which explains that performing the normalising also improves the data and hence produces stronger correlations. In brief, eliminating the weaker stimuli and normalising the data assists in refining the sentences, and so helps the search for the effective sentence type.

## 5.4 Discussion

This study focuses on (1) finding the effective sentence stimuli for measuring the cognitive processes of language processing, (2) exploring the different role of pauses in order to obtain information about the nature of language processing and (3) testing and confirming the findings of the pilot study in regard to IC and the Median. The present study (Experiment 1) improves upon the pilot study by recruiting a larger number of participants from a more focused population group (Malaysians). The experiment also introduced an independent English test, the ODT, as a means of gathering recent language performance data for the purpose of validating this method.



The discussion that follows first evaluates the findings from the pilot in order to confirm if in order to confirm whether these findings were effective, then assesses the effective sentence stimuli, and finally examines the many approaches of using pauses to measure the cognitive processes of language processing.

#### 5.4.1 Confirming the Pilot Findings

The findings in this study agree with the pilot in that they show that IC is suitable as a method of copying and that Median as a statistical measure is a suitable measure for cognitive processes in language processing, as the latter was able to provide significant correlation results within this experiment. The median of L2 pauses is found to be significantly strong in this study. This differs with the findings in the pilot, where L1 (letter) level correlations were stronger than L2 (word) level results, although both L1 and L2 had good correlations. In the pilot, it was discovered that one of the participants may have been dyslexic, as they showed relatively slow copying performance across all sentences, at L1 and L2 levels, compared with other participants. This may have made an impact on the effects on L2 correlations.

In brief, these findings support the approach also being used in the next experiment, in order to further corroborate them.

#### 5.4.2 Which Sentence Stimulus is Effective for Measuring Language?

The analysis found that the most effective sentence stimuli, based on the grouping by the significant value, labelled as the *best 6*, share general common characteristics: jumbling and ordinary sentences. These characteristics are believed to relate to sentence meaning, sentence structure, word spellings, and word knowledge and recognition. Two of the sentences in the *best 6* were garden path sentences, and this might not really count as a jumbling effect, these sentence types instead being generally processed as an ordinary sentence. Another sentence with a jumbling manipulation is that with the omission of spaces between words. This is treated as a jumbling condition because the omission of spaces between words increases the difficulty in distinguishing the words. There is a possibility that participants might recognise the wrong words because they are confused.

The decision to continue studying the jumbling effect is based on the majority of the sentence characteristic in the group (i.e. four sentences out of the best six contained jumbling conditions).

A detailed observation of the effects of garden path sentences showed that participants treat this sentence as a normal sentence. This observation was made by examining the pause distribution pattern across each letter and word in the sentence, which showed no long pauses. Long pauses are associated with the ambiguity point within the garden path sentence and may signify the unfamiliarity towards this type of sentence, so their absence is taken as an indication that participants are processing the sentence as a normal sentence with normal grammatical structure.

In brief, the findings indicate that the use of jumbling produces an effect in processing which can differentiate the nature of language processing. These stimuli have been shown to be capable of differentiating the competence levels of different participants. A competent participant, who has greater familiarity with the sentence or words, will be able to recall them easily from memory. This reduces the time spent searching for words and meanings, reduces processing and, hence produces shorter L2 (word) level pauses. By contrast, less competent participants, who are not familiar with such stimuli, require more processing time to digest the meaning of the whole sentence and hence produce longer pause lengths. This same approach to explaining the occurrences can be used for the rest of the sentence stimuli, with the conclusions differing depending on the condition of the sentence manipulation.

Familiarisation seems to play an important role in predicting how quickly a participant can copy the sentence stimuli presented. As explained in the literature (Chapter 2), familiarisation is also related to the ability to sight read, ability to recognise word shapes and the ability to process long words easily. As pause lengths represent the processing needed, the level of familiarity can be used to predict the pause lengths. For example, being competent in a language implies having ready access to an extended vocabulary database, as well as signifying that the recall and retrieval of information in order to complete a task do not need to refer to the long-term memory, but can just be accessed through the short-term memory. It is these processes of retrieving and recalling information related to the stimuli that determine pause lengths.

Feedback from the participants was gathered through short interviews carried out after the completion of the copying tasks. Participants agreed that some sentence stimuli were challenging. Competent participants find certain manipulations, such as the omission of spaces between words and the jumbling of letters in words, very confusing. Such types of manipulation disrupt the

visual recognition of words. Instead of being able to copy them fluently, the competent participant had to distinguish the words carefully while copying them in order to avoid mistakes. A good example of this is the 'jumbled proverb', which tests the participants' immediate response when copying begins. Even though competent participants might be familiar with the proverb, the jumbling of the words increases the processing time, because of the changes in the sentence structure. The fluency of copying the proverb is disturbed because of the different positioning of the words. Instead of copying fluently from memory, it is necessary to concentrate more on the stimulus in order to make sure it is copied correctly.

Such manipulation has definitely affected competent participants in terms of the need of increased processing time to process these manipulations. Less competent participants, on the other hand, are affected differently. Having limited vocabulary knowledge might make it more difficult for them to comprehend the sentence, so they are likely to produce long L2 (word) level pauses. If the language is largely unfamiliar, long sentence stimuli might appear as lists of words. This makes the copying task easier, but meaningless, because of the inability to build upon the meaning of the sentence based on the words. When a long word contains jumbled letters, less competent participants would be most likely to produce smaller chunks.

### **5.4.3 The Role of Pauses in Obtaining Information on Language Processing**

Pauses captured from the baseline copying were found useful as a means to normalise the data. Further findings in this experiment have shown that copying baselines (L1s) is strongly correlated to the L1s of sentence-copying. This also shows that applying L1 (letter) pauses from the sentence-copying has a similar effect to using the baselines (which focuses on the letter writing). This leads the way for the simplification of establishing baselines in designing the method of measuring the cognitive processes of language processing in this thesis. However, before proceeding to this final conclusion, it is important to understand what baseline tasks do and how this relates to the normalisation technique.

#### **5.4.3.1 The Baseline**

For the baseline copying or writing, the highly practised letters and words, not an indicator for measuring language, were chosen. They only focus on the fluency of handwriting production of letters and words, i.e. the speed of writing or writing performance. Therefore, baseline stimuli

were used as checks of the writing speed or of mistakes commonly occurring in the sentence stimuli copying processes. For those reasons, baselines do not correlate well with the ODT (Table 5.4). However, the correlation for 'upper-case name writing' is quite high ( $r(18)=-0.605$ ,  $p<.01$ ) which is unexpected and suggests the need for further investigation. As predicted, however, the baseline stimuli correlate significantly with most of the sentence stimuli, indicating a consistent writing performance. Findings indicated that the most effective baseline type is the one that involves 'lower-case name writing', as observed at both L1 and L2 pause levels.

Similar findings were observed when these correlations were investigated at each sentence level for L1 and L2 pause levels. The correlation values, however, vary across the sentences, and unexpected occurrences are revealed for sentences number 24 and 25 at L2 pause levels. The kind of manipulation used with these sentences (low frequency words, no spaces between words, combination of bigrams at the end-letters) might restrict the ability to distinguish individual words, which may be the reason for the weak correlations. The correlations, as can be seen in Figure 5.4 below, represent the occurrences of short L2 pauses at word level and long L1 pauses at letter level. An example of this can be seen in the graph below, which shows a comparison between the extreme cases of the most competent and least competent participants. Many long pauses occur at L1 (syllable or letter) level, represented by the dashed-line arrows and only a few long pauses at L2 (word) level, represented by the solid-line arrows. The fact that there is almost a flat graph-line during the beginning of the copying for the MC might suggest that either they are familiar with the words, which helps them copy fluently, or that they are processing the copying at syllable level and are very fluent at doing so. It could also be because of the ability to process a number of tasks simultaneously, referred to as 'parallel processing'. A detailed analysis of this can be found in Chapter 8.

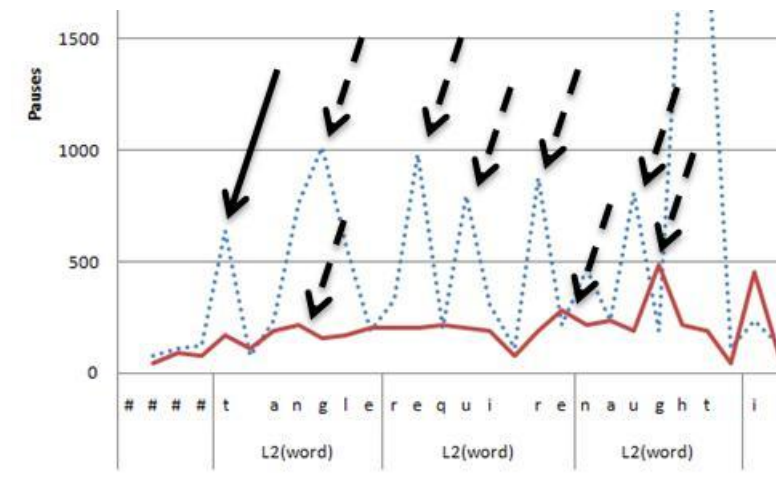


Figure 5.4 An example of pause distribution detail for sentence number 24, to observe the occurrences of long pauses that can be seen between the most competent (solid red line) and least competent (dotted blue line) participants

In short, participants who copy baseline stimuli with shorter pause lengths are most likely to produce shorter pauses when copying the sentence stimuli, and those who take longer copying the baselines are likely to take longer copying the sentence stimuli. Thus, the baselines are good measures of writing performance and using them is a good approach to identify issues that are unrelated to language processing, such as the individual writing speed. However, since the correlations between each baseline and the mean (average) of all sentences pauses show strong positive correlations (Table 5.4), the use of baselines might be unnecessary, as this information is also represented by the L1 (letter) level pauses captured from the sentence-copying. This may be a much simpler approach to measure the cognitive processes.

#### 5.4.3.2 The Normalisation

The application of normalisation has contributed significantly to the refinement of the sentence stimuli by compensating the pauses and by grouping the sentences according to the correlation values achieved. Normalisation introduced two new measures: *pause quotient* and *pause difference*. These are both aimed at correcting the pauses produced at both L2 and L1 pause levels. They are calculated either by dividing L2 pauses by L1 pauses (i.e. L2/L1) (*pause quotient*) or by calculating the difference between L2 and L1 (*pause difference*). Both measures were explored in order to identify which could better the original results of L2 Medians. However, both measures of normalising were able to produce strong significant correlations.

Despite this, it is essential to choose the measure that allows further exploration of normalisation in the coming chapters. The pause for writing the first letter at the beginning of a word includes both the processes for preparing for the whole word and the process for writing that particular letter. It would therefore be more theoretically coherent to assume that taking away the time needed for the sub-process (letter level pause) is a better method than dividing the total time taken for the word by the number of letters in a word, which would hinder the recognition of different pause distributions across that particular word. This suggests that *pause difference* is a logical choice for further exploration of normalisation.

#### 5.4.4 The Implications of the MoC

The conclusion that can be made in relation to the model can be seen from the six sentences with the strongest correlations. Close analysis of the correlations at L1 letter level against ODT for all participants and each sentence shows weak correlations for pauses being longer at letter level within a word, but instead shows a clear strong correlation for long pauses to occur at L2 word level.

What this means for the model is that the route to processing words and letters does exist and that long pauses produced may be due to semantic processing. For example, the garden path sentences were treated like normal sentences: words were copied individually or in groups in the process to build up the meaning, but long pauses were not found at the ambiguity point, showing that the participants were not able to build up the actual meaning of the garden path sentence, as required. This may also be due to unfamiliarity with garden path sentences by participants of English as their second language; they have difficulty understanding of a garden path sentence because there are two different meanings in a sentence. Thus, long pauses were found at L2 word level just like for a normal sentence, as this is what they are familiar with.

There are different possible routes that may be used in a sentence. This may be for many different reasons, such as familiarity level of the words to be copied, or the presentation of the words or sentence to the participants, such as the jumbled conditions. For example, one might find it easy to copy a group of high frequency words together as one big chunk, e.g. 'there is a car', because these words are familiar and hence do not need semantic processing. However, there may be times in the sentence where one would take a different route in order to copy, because a particular word may be of low frequency, lengthy or not familiar, e.g. 'indespicable'.

The copying of the word 'indespicable' may take a repeated route of copying a group of letters or syllables, until the word is complete.

It is also possible that someone would use just one route of copying regardless of other factors, because of the familiarity of that particular route as a method to copy fluently. For example, when a person is copying a sentence where letters in all of the words are jumbled, they might decide to copy by grouping letters together throughout the copying process as this is the easiest way for them to copy as quickly as possible while trying to build up the sentence meaning. Sometimes, in the case of unfamiliar or difficult words, the copier may not have any other choice but to copy letter by letter, perhaps because that is the only way to copy as quickly as possible, in order to complete the task.

As a conclusion, the model assists this research in understanding the different routes someone might take in order to process the copying, either at word level or letter level. Furthermore, the model also assists in understanding what constitutes a pause. A pause may be long or short, and this length depends on the routes taken, as shown in the model. As explained, a sentence may include different pause lengths, depending on a person's familiarity with the words presented. Although the model is able to explain this, it is not able to show a generalisation of neither copying pattern nor definite pause lengths for everyone copying the same sentence, because each individual will take a unique approach.

## 5.5 Conclusion

This experiment has investigated the effective sentence stimuli for measuring the cognitive processes in language processing. The study indicates that the sentence types which best differentiate between the participants' nature of language processing are ordinary sentences and sentences which use the element of jumbling, suggesting a need for further investigation focusing on these elements.

This experiment was limited in that the sentence stimuli that were applied were not systematically randomised. In other words, all participants copied them in the same sequence, which might cause the element of fatigue to affect the results. However, the data and results are

clear of any fatigue effects and the lack of systematic control over the stimuli randomisation is therefore considered inconsequential at this stage.

The first intention of this experiment was to measure language in general, but the use of ODT has highlighted that the sentence-copying is able to measure grammatical knowledge, which is a component of language. This suggests that sentence-copying is able to measure specific components of language. The correlation findings with ODT indicate the ability of the sentence-copying to measure grammar competence. This indirectly suggests the possibility of exploring the measuring of cognitive processes components in language processing. Hence, in the next experiment, language components, such as word knowledge, will be considered.

The 'normalisation' technique has been proven useful in correcting for individual writing speeds, as it increases the correlation values between the L2 pauses and the ODT scores. Further, the observations made between the L1 baseline and L1 sentences demonstrate a strong correlation, which indicates that applying the L1 sentences is sufficient for the purpose of normalisation. Moreover, the normalisation technique has introduced two new calculations, *pause quotient* and *pause difference*, which may be applicable in later experiments.

Above all, there is also the need to understand the processes involved in copying that directly affect the different pause lengths, such as how a specific jumbling condition affects the different pause lengths produced by different nature of language processing. On this basis, it is argued that a theoretical framework (explained in Chapter 4) is useful for presenting a general overview of the predicted processes that are involved in the activity of sentence-copying.

To conclude, this experiment has shown that the suitable sentence type for differentiating the nature of language processing involves the element of *jumbling* and ordinary sentences. In the next experiment, the effect of these conditions is explored further. In doing so, however, the design of the sentence stimuli must become more restricted, so guidelines to 'jumbling' will be introduced.



## CHAPTER 6    Experiment 2: Investigating a Suitable Jumbling Condition to Assess Components of Language Processing

### 6.1 Introduction

This chapter aims to investigate the effects of using jumbling as a sentence manipulation. This is being investigated following data from the previous experiment, where sentences which obtained strong correlations between the sentence-copying and ODT (an independent language test) consists of ordinary sentences and sentences with jumbling conditions such as words and letters jumbled. In order to assess the nature of language processing with measures of language components such as grammatical and vocabulary knowledge, the study thus used four kinds of jumbling: (1) *no jumbling* (normal sentences), (2) *words jumbled*, (3) *letters jumbled*, and (4) *both words and letters jumbled*; to attempt to differentiate the cognitive processing levels within the group of participants.

These four types of jumbling were considered to be sufficient for sentence level manipulations as they can be controlled in such a way that will affect the grammar structure (*words jumbled*) or vocabulary knowledge (*letters jumbled*). Instead of applying a random jumbling concept, the jumbling was performed based on certain rules, which will be explained in great detail in this chapter. The rules were created in order to ensure a consistent approach to manipulating the sentences, and consistent sentence stimuli design. This approach not only focuses on testing grammar knowledge, but also extends to the investigation of word knowledge. There is therefore a need for a vocabulary test. The Vocabulary Size Test (VST) by Nation (2007) was chosen.

In relation to each jumbling condition, it is predicted that the *no jumbling* condition will test participants' level of comprehension at sentence level; the *word jumbled* condition will test grammar knowledge; the *letters jumbled* condition will test vocabulary knowledge and the *both jumbled* condition will be the most difficult of all four. With the *both jumbled* condition, there is a possibility that participants will struggle to copy and therefore give up because the sentence may appear meaningless. Using different levels of complexities degrades the normal process of copying, affecting the size of chunks and pause lengths, hence allowing the assessment of

different levels of language processing among individuals. The least competent and the most competent participants are labelled LC and MC respectively.

Academic sentences were chosen from various journals in the arts and sciences in order to be able to control the level of difficulty by considering aspects such as the familiarity of technical terms (i.e. word frequencies), equivalent distribution of different word lengths, and the length of the sentence. This approach is more focused compared to the previous experiment, because the range of sentence stimuli is less varied.

This experiment also aims to test the use of normalisation as a means to improve the measure by attempting to compensate for individual differences in writing speed, as findings from the previous experiment found normalisation to be useful. However, instead of using both *pause difference* and *pause quotient* in calculating the normalisation, this experiment will only use *pause difference*. The reason for choosing *pause difference* is explained in Section 5.4.3.2. It is hoped that performing normalisation will improve quality of the data captured.

As a variation to the recruitment of the population group in Experiment 1, this experiment employed 20 Spanish participants with ESL. It was considered important to test the approach on a different population. However, both populations must use the same script in their first language. This avoids issues with unfamiliarity of letter-writing that might additionally affect the sentence-copying processes.

Extending the aims for this experiment, this chapter will first look at the sentence stimuli design.

## 6.2 The Sentence Stimuli Design

This section will describe the sentence stimuli used in this experiment. First, the rationale for using the jumbling condition is provided. Then each jumbling type that will be used in this experiment is examined. Next, the section will describe the design of stimuli, including an explanation of how the structure of the sentence stimuli was created. Finally, the simulation of the sentence stimuli is presented by making predictions of the chunking processes that might occur for the two extreme cases of a *most competent* (MC) and *least competent* (LC) participant.

### 6.2.1 Why Jumbling Conditions?

The jumbling conditions that are focused on in this experiment involve the jumbling of words in a sentence and the jumbling of letters in a word. The use of the jumbling condition is motivated by the strong correlations obtained in Experiment 1 between the sentences and ODT. The two types of sentences used, which share similar characteristics to jumbling conditions are (1) words of proverbs jumbled by combining two proverbs and (2) letters jumbled in words. The garden path sentences however, were treated as ordinary sentences, as thorough examination of the pause distributions has not been able to provide evidence of any long pauses occurring at the ambiguous point of a garden path sentence. Sentence where the spaces between words were omitted, on the other hand, will not be focused on as it is predicted that the difficulty in distinguishing the words (because of the spaces) becomes redundant when words and letters are jumbled.

As a result, four types of jumbling were used: *no jumbling*, *words jumbled*, *letters jumbled* and *both jumbled*. The level of difficulty ranges from *no jumbling*, being the easiest to process, to *both jumbled* being the most difficult. *No jumbling* is considered the easiest because it contains correct grammatical structure and correct word presentation, so participants can use their understanding of the sentence meaning to help them process it. Sentences are more difficult to process when the words are jumbled and the grammatical structure is disrupted, because it is harder to make out the meaning of the whole sentence. When letters in a word are jumbled, the disruption increases the processing time because of the need to deal with the letters in an unexpected order, also increasing the difficulty of the comprehension of the whole sentence. Accordingly, when both words and letters are jumbled, it makes the process of trying to comprehend the sentence doubly difficult. Detailed descriptions of each jumbling type are provided in the next section.

### 6.2.2 Levels of Jumbling

The four types of jumbling are based on a 2x2 factorial design, illustrated in figure 6.1. For every sentence, four different kinds of manipulation were constructed, from the combinations of normal (N) or jumbled (J) conditions, at the word (W) or letter (L) level. Throughout this thesis, the jumbling condition is referred to as *no jumbling (NWNL)*, *words jumbled (JWNL)*, *letters jumbled (NWJL)* and *both jumbled (JWJL)*. In the subsequent subsections, each of the jumbling

types is described and predictions made about the process of copying following the theoretical MoC.

		WORDS (W)	
		Normal (N)	Jumbled (J)
LETTERS (L)	Normal (N)	NWNL	JWNL
		no jumbling	words jumbled
	Jumbled (J)	NWJL	JWJL
		letters jumbled	both jumbled

Figure 6.1 2x2 Factorial Design of the Experiment

#### 6.2.2.1 No Jumbling Condition

*No jumbling* has correct grammatical structure and correct letter positions, preserving the sentence's grammatical structure and word presentation. Although this may be thought to be the easiest to comprehend of the four conditions, the sentence can be difficult if it contains low frequency words, such as technical terms used in specific academic fields, which are likely to be unfamiliar to the participant.

It is predicted that participants who are of low competence and unfamiliar with the words in the sentence will (1) use small chunks at word and syllable level, and (2) require higher PLP values, and thus produce many long L2 and L1 pauses. Participants who are of high competence and familiar with more of the words in the sentence will (1) have a larger chunk size at word group level, and (2) require lower PLP values, and thus produce fewer long L2 and L1 pauses.

#### 6.2.2.2 Words Jumbled Condition

The *words jumbled* condition has a disordered grammatical structure, because the words have been moved around, but the word spellings are still preserved. In order to systematically control the word jumbling manipulations, each sentence was divided in half and alternating words were taken from the two halves, with the first word of the second half used as the first word of the jumbled sentence. The aim is to eliminate the original meaning of the sentence.

It is predicted that participants who are of low competence and unfamiliar with the language might not realise that the words are jumbled, and may treat the sentence as a list of disconnected words. As a result, they will have a small chunk size at word and syllable level and show higher PLP values, thus producing many long L2 and L1 pauses. Participants who are of high competence and familiar with the language might realise that the grammar structure has been disturbed and will therefore still be able to produce large chunk sizes at word group level, although possibly smaller than with *no jumbling* type, and would demand lower PLP values, consequently producing a number of long L2 and L1 pauses.

#### 6.2.2.3 *Letters Jumbled Condition*

The *letters jumbled* condition has correct grammatical structure, but the letters in the words are jumbled. This jumbling type manipulates the spellings, where the letter position within the word is transposed in a controlled manner, the aim being to test word recognition competence. The transposition of letters is controlled in such a way to avoid copying at a non-meaningful letter-by-letter level. It is, however, intended that participants would at the very least perform syllable level chunking, as it is assumed that adult participants already have some phonemic and graphemic knowledge (Chapter 2) that would enable the process of copying groups of letter within the whole word level.

It is predicted that participants who are of low competence and unfamiliar with certain words in the sentences might find it difficult to recognise them because of the jumbled letters. As a result, they might produce many small chunk sizes at word and syllable level and show higher PLP values, thus producing many long L2 and L1 pauses. Participants who are of high competence and familiar with the words in the sentence might still be able to grasp the words by the *word shape* (Chapter 2), and still chunk at word group level, but only with short word lengths. Such participants may also be prone to producing many small chunks at word or syllable level because of the disordered spelling, thus showing higher PLP values because of the difficulty, and accordingly produce many long L2 and L1 pauses.

In order to control this jumbling condition, each word is changed according to the rules listed below. The changes are, however, controlled so that the shape of the word is still preserved, to enable the process of word recognition.

- Rule #1: First, preserve the end letters (the first and the last letter of each word).
- Rule #2: Preserve the beginning bigrams (beginning bigrams are the first two letters in a word, which include the first end-letter).
- Rule #3: Transposing letters must first use the consonants (C) alone and must be done across morphemes: C and C.
- Rule #4: If this is impossible, vowels (V) are then allowed for an exchange. These rules, however, must be followed: C and V or V and C; then only V and V.
- Rule #5: For 4 letter words, only swap the internal letters (rules number 2, 3, 4 are void).
- Rule #6: For 5 letter words, rules number 1 and 2 apply. Only swap the remainder 2 letters.
- Rule #7: 1, 2 and 3 letter words remain the same.

Below are examples of letter jumbling following the rules.

For words containing more than 5 letters, e.g. 'preserve':

- Rule #1:* p \_ \_ \_ \_ e ; the end-letters are preserved.
- Rule #2:* p r \_ \_ \_ e ; the beginning bigrams are preserved.
- Rules #3 & #4:* p r r s e e v e d ; the morphemes are 'pre' and 'served'. Rule #4 dictates a V and C swap; hence, the 'e' and 'r' are exchanged, as boldfaced in the word.

For words containing only 4 letters, e.g. 'path':

- Rule #5:* p \_ \_ h ; Only swap the middle letters.
- = p **t** **a** h

For words containing only 5 letters, e.g. 'there':

- Rule #6:* t h \_ \_ e ; Rules #1 and #2 apply. Only swap the remainder letters.
- = t h **r** **e** e

#### 6.2.2.4 *Both Jumbled Condition*

The *both jumbled* condition has both the grammatical structure and word spellings mixed up. This jumbling type is trickier than the rest as it combines both word and letter jumbling. With this kind of sentence, the study assumed that it is difficult to recognise the words, making it complicated to connect each of the words into a meaningful construct and thus impossible to comprehend the meaning of the whole sentence.

It is predicted that participants who are low of competence and unfamiliar with the words in the sentence will produce many small chunk sizes at word and syllable level, and show the highest PLP values, thus producing many long L2 and L1 pauses. Participants who are of high competence and familiar with the words in the sentence might not be able to produce larger chunk sizes at word group level, but instead might produce many small chunks at word and syllable level, similar to the less competent participants. Competent participants might show higher PLP values because of their difficulty in processing the words and sentence, thus producing many long L2 and L1 pauses. Although the manipulation has rendered the condition of a competent participant similar to that of the less competent, they may still engage different copying processes based on their knowledge level. As a result, it is predicted that this type of manipulation will be the most difficult to copy and will result in many long L2 and L1 pauses, the longest overall pauses out of the four jumbling conditions.

#### 6.2.3 *The Design of Stimuli*

Twelve sentences were taken from 12 different research journals from different fields, including Art History, Biology, Business, Mathematics, and Linguistics. The difficulties of each sentence varies in terms of the technical terms used (i.e. word frequencies). However, these word frequencies are controlled by ensuring that each sentence has approximately 50% high frequency words (0–1000 frequency level) and 50% low frequency words (beyond the 1000 frequency word level).

To measuring the word frequencies, Laufer and Nation's Lexical Frequency Profiler was used, an online computer program which divides words into first and second thousand levels, academic words, and the remainder, or 'offlist' (Laufer & Nation, 1995; 2006). The offlist (i.e. remainder) comprises words that do not belong to the first and second thousand levels and are not academic words. The program, known as Vocabulary Profile (VP) was first developed in 1995 and was last

upgraded in January 2006. This thesis refers to the high frequency words as those within the first thousand frequency level, and low frequency words as being the second thousand frequency level, the academic words and the offlist. VP has been used by researchers such as Laufer and Nation (1995, 2006) as a research instrument to study vocabulary; Meara (1993) as a means to evaluate an English course; and Meara and Fitzpatrick (2000) to assess productive vocabulary.

#### 6.2.4 Making Simulations with the Jumbling Conditions

The aim of this section is to predict the relative performance of participants with different levels of competence in the different types of jumbling. The predictions are made based on the theoretical Model of Copying (MoC), as introduced in Chapter 4. The simulations were done by predicting the processing routes that would be taken in order to process the copying. For the purpose of making predictions, two participants with extreme levels of competence are imagined, one with high competence, MC (most competent) and one with low competence, LC (least competent). The predictions were made to forecast the pause lengths that might take place during the process of copying.

In order to make the predictions, a point of reference is needed as a basis for the process. Since sentences consist of words, and words can be easily identified with L2 pauses, it was decided to use words as a reference to make the predictions. Moreover, the MoC itself is largely based on previous literature that focused on word recognition. Words were therefore classified into eight categories, as explained in the next section. These words are used as a means to predict where chunking will occur in a sentence.

To make these predictions, *Processing Load Prediction* (PLP) values are assigned for the chunking processes that are expected to occur. These processes depend on the levels of competence. Detailed explanations of the PLP can be found in Chapter 4, Section 4.3.3, where a full description for each PLP name is given. The predictions made with regard to PLPs and levels of competence are presented in Table 6.2 (Section 6.2.4.2).

##### 6.2.4.1 Word Classifications as a Basis to Making Predictions

The classifications of word characteristics defined in Section 4.3.3 are presented again here, for ease of reference (Table 6.1). The predicted chunking processes at word level will use the eight word characteristics shown in Table 6.1. The eight word characteristics are based on three



criteria: (1) word length, (2) word frequency and (3) familiarity level of words, hence the naming convention of characteristics in the form of ‘length-frequency-familiarity.’ Each of these word characteristics is numbered as in Table 6.1 used as a reference for making predictions in Table 6.2.

	SHORT		LONG	
	Familiar	Unfamiliar	Familiar	Unfamiliar
High Frequency	<i>Sho-Hi-Fam (1)</i>	<i>Sho-Hi-Unfam (2)</i>	<i>Lo-Hi-Fam (5)</i>	<i>Lo-Hi-Unfam (6)</i>
Low Frequency	<i>Sho-Low-Fam (3)</i>	<i>Sho-Low-Unfam (4)</i>	<i>Lo-Low-Fam (7)</i>	<i>Lo-Low-Unfam (8)</i>

Table 6.1 Word Characteristics

In Table 6.1, word length is defined as short when the word consists of 4 letters or fewer, otherwise it is long. Word frequency is defined as high when the word is in the first thousand frequency level (Laufer & Nation, 2006), otherwise it is defined as low. Word familiarity was defined and classified based on the author’s own discretion, which is as designed in the agreement of MoC. The author’s decision is based on predictions made with the MoC of how the particular word may be processed by the most competent or least competent participants.

As an example, short, high frequency and unfamiliar words will be more common for MC; the same word may be familiar to LC. In the case of the words ‘best’ and ‘conception’, ‘best’ could be easily classified as word characteristic (1) of Sho-Hi-Fam for both MC and LC, while ‘conception’, which is more than a four letter word, would belong to the *Long* column, but may be (5) Lo-Hi-Fam for MC and (8) Lo-Low-Unfam for LC. Classifying word characteristics for individual’s levels of competence is not a straightforward process, and depends on their language knowledge.

Level of Processing	No Jumbling		Words Jumbled		Letters Jumbled		Both Jumbled	
	MC	LC	MC	LC	MC	LC	MC	LC
<b>Word Group Level Chunking</b>	Likely to happen - Many large chunk sizes	Possible to occur - Few large chunk size	Likely to happen - Many large chunk size	Possible to occur - Few large chunk sizes	Possible but unlikely to occur - Few large chunk sizes	Possible but very unlikely to occur - Very few large chunk sizes	Possible but very unlikely to occur - Very few large chunk sizes	Possible but very unlikely to occur - Very few large chunk sizes
<b>Predictions</b>	For all word types. Number of words in a group of chunk could vary between 2 and 4, depending on the length, but the maximum would be 4	For Word Types 1,3,5	For all word types. Number of words in a group will be lesser than in SenType 'a'	For Word Types 1,3,5	Possible only for copying words that contain 4 letters or fewer, if they appear next to each other	Possible only for copying words that contain 4 letters or fewer, if they appear next to each other	Possible only for copying words that contain 4 letters or fewer, if they appear next to each other	Possible only for copying words that contain 4 letters or fewer, if they appear next to each other
<b>PLP Values Predicted</b>	Word Types 1,2,3,5,7: <b>PLP1</b> ; Word Types 4,6,8: <b>PLP2</b>	Word Type 1,3,5: <b>PLP1</b>	Word Types 1,2,3,5,7: <b>PLP1</b> ; Word Types 4,6,8: <b>PLP2</b>	Word Types 1,3,5: <b>PLP1</b>	Word Type 1*: <b>PLP1</b>	Word Type 1*: <b>PLP1</b>	Word Type 1: <b>PLP1</b>	Word Type 1*: <b>PLP1</b>
<b>L2 Pause Predictions</b>	Many long L2 pauses (however shorter than LC)	Few long L2 pauses (which is longer than MC)	Many long L2 pauses (but a lower number of words compared to No Jumbling for MC)	Few L2 pauses (only for high frequency words that contain 4 letters or fewer, appearing side by side)	Few L2 pauses (only for high frequency words that contain 4 letters or fewer, appearing side by side)	Few L2 pauses (only for high frequency words that contain 4 letters or fewer, appearing side by side)	Few L2 pauses (only for high frequency words that contain 4 letters or fewer, appearing side by side)	Few L2 pauses (only for high frequency words that contain 4 letters or fewer, appearing side by side)
<b>Word Level Chunking</b>	Possible	Likely	Likely	Likely	Likely	Possible	Likely	Possible
<b>Predictions</b>		For Word Types 1,2,3,4,5		For Word Types 1,2,3,4,5	Possible for words that contain more than 4 letters. For Word Types 1,2,3,4	Possible for words that contain 4 letters or fewer. For Word Types 1,2,3	Possible for words that contain more than 4 letters. For Word Types 2,3,4	Possible for words that contain 4 letters or fewer. For Word Types 1,2,3
<b>PLP Values Predicted</b>		Word Types 1,3,5: <b>PLP3</b> ; Word Types 2,4: <b>PLP4</b>		Word Types 1,3,5: <b>PLP3</b> ; Word Types 2,4: <b>PLP4</b>	Word Types 1**,3: <b>PLP3</b> ; Word Types 2,4: <b>PLP4</b>	Word Types 1**,3: <b>PLP3</b> ; Word Type 2: <b>PLP4</b>	Word Type 3: <b>PLP3</b> ; Word Types 2,4: <b>PLP4</b>	Word Type 1*: <b>PLP3</b> ; Word Types 2: <b>PLP4</b>
<b>L2 Pause Predictions</b>		Many long L2 pauses		Many long L2 pauses	Many long L2 pauses (shorter than LC)	Many long L2 pauses (longer than MC)	Many long L2 pauses (but overall pauses are shorter than LC)	Many long L2 pauses (but overall pauses are longer than MC)
<b>Syllable / Letter Level Chunking</b>	Possible but unlikely unless unfamiliar with spelling	Very likely to happen for long words	Possible but unlikely unless unfamiliar with spelling	Very likely to happen for long words	Very likely to happen for jumbled words that have more than 4 letters	Very likely to happen for long words (4 letters and above)	Very likely to happen for jumbled words that have more than 4 letters	Very likely to happen for long words (4 letters and above)
<b>Predictions</b>		For Word Types 5,6,7,8		For Word Types 5,6,7,8	For Word Types 5,6,7,8	For Word Type 4, 5,6,7,8	For Word Types 5,6,7,8	For all word type
<b>PLP Values Predicted</b>		Word Types 5,7: <b>PLP6</b> ; Word Types 6,8: <b>PLP7</b> ; <b>PLP5</b> for within a word		Word Types 5,7: <b>PLP6</b> ; Word Types 6,8: <b>PLP7</b> ; <b>PLP5</b> for within a word	Word Types 5,7: <b>PLP6</b> ; Word Types 6,8: <b>PLP7</b> ; <b>PLP5</b> for within a word	Word Types 5,7: <b>PLP6</b> ; Word Types 4,6,8: <b>PLP7</b> ; <b>PLP5</b> for within a word	Word Types 5,7: <b>PLP6</b> ; Word Types 6,8: <b>PLP7</b> ; <b>PLP5</b> for within a word	Word Type 1**,3,5,7: <b>PLP6</b> ; Word Types 2,4,6,8: <b>PLP7</b> ; <b>PLP5</b> for within a word
<b>L1 Pause Predictions</b>					Few L1 pauses	Many long L1 pauses (with more L1s compared to MC)	Many long L1 pauses (but overall pauses are shorter than LC)	Many long L1 pauses (but overall pauses are longer than MC)

1\* = words with <=4 letters; 1\*\* = words with >4 letters

**Table 6.2 Predictions of possible processes occurring during copying for the extreme conditions of Most Competent and Least Competent participant**

Table 6.2 presents the predictions made for each level of processing, from word group level to letter level, to show how chunking size varies for each jumbling type compared between the two extreme cases of MC and LC participants. The predictions are based on the *word characteristics* as in Table 6.1, and also suggest the *PLP values* and *the Pauses predicted* accordingly. To recap from Chapter 4, processing levels include of processing at word group, word, syllable and letter level. These levels of processing are measured by *pause levels* L2 (word) and L1 (letter). Figure 4.4 in

Chapter 4 shows the relation between levels of processing and the pause level values. However, as the chunk an individual uses depends on their competence level, the prediction of where chunks (and hence long pauses) will occur is difficult. Thus, a suitable approach would be to make predictions based on the word characteristics, as proposed, in order to examine the effects of copying the four jumbling types.

In order to give examples of how predictions are made using Table 6.2, the following section will provide explanations for each of the jumbling types.

#### **6.2.4.2 Simulation of the Jumbling Conditions**

This section will describe the simulations based on the four jumbling conditions as described in the previous sections. The predictions are made with the guidance of Figure 4.9 (Section 4.4.4) and the MoC, where the PLP values are used to predict the possible pause values occurring at each chunk. For each jumbling condition, the following are described: (1) the number of chunks occurring at word group and word level, (2) the number of chunks occurring at group, word and syllable level, (3) the median of PLPs, excluding letters, and (4) the results of the PLP values plotted on a graph, as a conclusion to the predictions.

#### **6.2.4.3 Predictions for No Jumbling Condition (NWNL)**

This sentence is used as an example for making predictions using the MoC for MC and LC participants:

*“It normally involves having access to a secure site on the internet where a graded series of lessons are available.”*

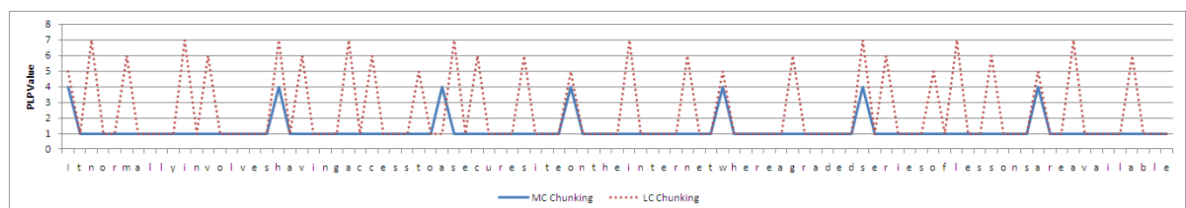
Participants are expected to copy this fluently. With this kind of sentence, participants' ability to comprehend the sentence meaning will be tested. What differentiates the levels of competence is the familiarity of the words used in the sentence, which determines the way participants chunk and the MoC PLP route that are taken, affecting the pause lengths. The example sentence uses words that can be easily comprehended and may be of high familiarity, and has the correct grammatical structure. This means that the routes taken in the MoC are straightforward, comprising short overall (L2 and L1) pause lengths, and are likely to be among the shortest total route. It is expected that MC will have a large chunk size compared to LC. Large chunks size may

consist of several words familiar to the participant. Further possibilities of chunking can be seen in Table 6.2.

No Jumbling	i t n o r m a l l y i n v o l v e s h a v i n g a c c e s s t o a s e c u r e s i t e o n t h														
MC Chunking	4 4 4														
Group / Word / Syllable Level															
Letter Level	4 1 1 1 1 1 1 1 1 1 1 1 1 1 1														
No Jumbling	i t n o r m a l l y i n v o l v e s h a v i n g a c c e s s t o a s e c u r e s i t e o n t h														
LC Chunking	5 7 6 7 6 7 6 5 1 7 6														
Group / Word / Syllable Level	5 7 6 7 6 7 6 5 1 7 6 6 5 1														
Letter Level	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1														
No Jumbling	e i n t e r n e t w h e r e a g r a d e d s e r i e s o f l e s s o n s a r e a v a i l a b l e														
MC Chunking	4 4 4														
Group / Word / Syllable Level	1 4 1 1														
Letter Level	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1														
No Jumbling	e i n t e r n e t w h e r e a g r a d e d s e r i e s o f l e s s o n s a r e a v a i l a b l e														
LC Chunking	7 6 5 6 7 6 5 7 6 5 7 6														
Group / Word / Syllable Level	7 6 5 7 6 5 7 6 5 7 6														
Letter Level	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1														

**Figure 6.2 Sentence example for no jumbling condition**

Figure 6.2 analyses the copying of the sample sentence by MC and LC. At word group level and at word level, MC produces 7 large chunks while LC produces 17 smaller chunks. At group, word and syllable level, MC still produces 7 large chunks, but LC has 26 smaller chunks (includes the grey vertical border). MC performs chunking at word group level, but LC mostly processes word by word, or even at syllable level. When MC processes a chunk that contains a number of words, the whole chunk is treated as one unit, and the PLP value is assigned to the first letter of the unit. The value of '1' is then given to each of the other words in the unit, in the same way that a value of '1' is given to each letter in a long familiar word. Overall, MC produces a median PLP of 4, and LC produces a median PLP of 6, ignoring the letter level PLP values as the focus is only on the chunking processes. LC has higher processing costs than MC and so is predicted to have longer pause lengths.



**Figure 6.3 PLP values for a no jumbling condition sentence**

Figure 6.3 shows the PLP values for both MC and LC. The PLP values are on the y-axis, and the x-axis displays the successive letters of the sentence. The solid line represents MC, while the dotted line represents LC. It can be seen that MC has fewer chunks than LC and that MC's PLP values are shorter than LC's. Thus, it is predicted that MC will have pauses that tend to be shorter than LC and occur less frequently.

#### 6.2.4.4 Predictions for Words Jumbled Condition (JWNL)

In this condition, the grammatical structure is manipulated by jumbling the words, so the previous same sentence now becomes:

*“the it internet normally where involves a having graded access series to of a lessons secure are site available on”*

With a disturbed grammatical structure, participants are not expected to copy as fluently as with the previous sentence. Participants will probably treat it as a disconnected list of words without an overall meaning, although there might be some possibility of a highly competent language user being able to capture the overall sentence meaning. Overall, what differentiates the levels of competence is the familiarity of the words used in the sentence, which determines chunk size and thus pause lengths. It is expected that MC would be able to produce a larger chunk size, by grouping unrelated words together, than LC, who would treat the sentence as a list of words.

Words Jumbled	t	h	e	i	t	i	n	t	e	r	n	e	t	n	o	r	m	a	i	l	y	w	h	e	r	e	i	n	v	o	l	v	e	s	a	h	a	v	i	n	g	g	r	a	d	e	d	a																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
MC Chunking	4														4													4																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
Group / Word / Syllable Level	4			1			1								4							1					4								1	1						4					1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 6.4 Sentence example for words jumbled condition

Figure 6.4 analyses the copying of the given sentence and Figure 6.5 shows the resulting PLP values. At word group level and word level chunking, MC produces 8 large chunks while LC produces 16 smaller chunks. At group, word and syllable level chunking, MC still produces 8 large chunks, but LC has 25 smaller chunks (includes the grey vertical border). MC is able to chunk

groups of words similar to in the *no jumbling* condition. LC is also able to chunk groups of words, but only when the words are of high frequency. They are mostly processed by LC at word level, or even at syllable level for unfamiliar words. The rest of the chunking processes are similar to the explanation given in the *no jumbling* condition. Overall, MC produces median PLP values of 4, and LC produces median PLP values of 6.

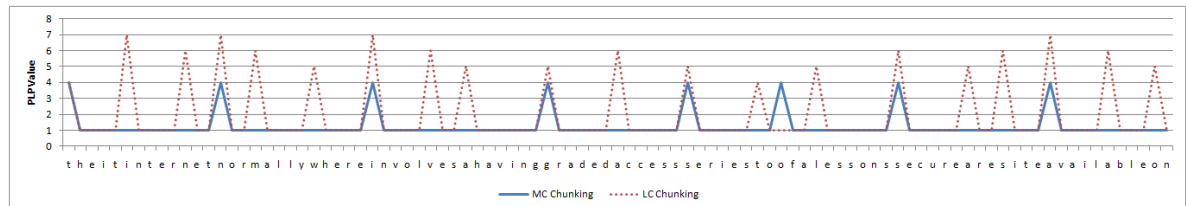


Figure 6.5 PLP values for a words jumbled condition sentence

Figure 6.5 shows that MC has fewer chunks than LC and that MC's PLP values are shorter than LC. Both Figure 6.3 (PLP values for a *no jumbling* condition sentence) and Figure 6.5 above provide chunking numbers and patterns which demonstrate that both jumbling conditions have the potential to differentiate the levels of cognitive processes. In brief, it is predicted that, in both conditions, MC will have pauses that tend to be shorter than LC with less frequent chunking.

#### 6.2.4.5 Predictions for Letters Jumbled Condition (NWJL)

This time, the letters within the words are jumbled:

*"It nolmarly inlovves hanivg acsecs to a seruce stie on the innertet whree a gredad sereis of lensoss are avaiballe"*

For the sentence above, it is predicted that both MC and LC will spend a long time copying, but for different reasons, mobilising different processes. Participants are most likely to focus on word level chunking. For all competence levels, the jumbled spelling interferes with the process of copying. Because of their greater familiarity with the proper spellings, MC might automatically correct the spelling, subconsciously. Longer pauses are therefore associated with the MC because they cannot rely upon remembered spelling. The overall process may be further slowed down by MC trying to connect the words to understand the sentence. LC, on the other hand, is unable to recognise the actual words at all, so will copy at syllable level. Therefore, the LC would produce many small chunks at syllable level, but will require more processing time, as they will take a

longer processing route to perform the segmentation of the words, and hence have long pauses in general. In conclusion, large processing costs are predicted in both cases.

Letters Jumbled	i	t	n	o	i	m	a	r	i	y	i	n	i	o	v	v	e	s	h	a	n	i	v	g	a	c	s	e	c	s	t	o	a	s	e	r	u	c	e	s	t	i	e	o	n	t	h		
MC Chunking	5		8				6					8		6					8		6				8		6				4				8		6			6			4						
Group of Words / Word / Syllable	5		8				6					8		6					8		6				8		6				4		1		8		6			6		1	1	1		4		1	
Letter	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Letters Jumbled	i	t	n	o	i	m	a	r	i	y	i	n	i	o	v	v	e	s	h	a	n	i	v	g	a	c	s	e	c	s	t	o	a	s	e	r	u	c	e	s	t	i	e	o	n	t	h		
LC Chunking	5		7			6		6				7		6		6			7		6				7		6				4			7		6			7		6		4						
Group of Words / Word / Syllable	5		7			6		6				7		6		6			7		6				7		6				4		1		7		6			7		6		4		1			
Letter	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Letters Jumbled	e	i	n	n	e	r	t	e	t	w	h	r	e	e	a	g	r	e	d	a	d	s	e	r	e	i	s	o	f	i	e	n	s	o	s	s	a	r	e	a	v	a	i	b	a	i	e		
MC Chunking		8					6			5					5		8		6				8		6			5		8		6			5			8		6				6					
Group of Words / Word / Syllable		8					6			5					5		8		6				8		6			5		8		6			5			8		6				6					
Letter	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Letters Jumbled	e	i	n	n	e	r	t	e	t	w	h	r	e	e	a	g	r	e	d	a	d	s	e	r	e	i	s	o	f	i	e	n	s	o	s	s	a	r	e	a	v	a	i	b	a	i	e		
LC Chunking		7			6		6			7		6		5		7		6				7		6		5		7		6		6		5			7		6		6		6						
Group of Words / Word / Syllable		7			6		6			7		6		5		7		6				7		6		5		7		6		6		5			7		6		6		6						
Letter	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 6.6 Sentence example for a letters jumbled condition

Figure 6.6 above analyses the possible copying of the sentence given, and Figure 6.7 shows the overall outcome. At word group level and word level chunking, both MC and LC produce 18 chunks, mostly at word level. At group, word and syllable level chunking, MC produces 28 small chunks and LC produces 35 smaller chunks (includes the grey vertical border). Both MC and LC demonstrate syllable level chunking. The increase in the number of chunks and the decrease in chunking size show the difficulty in copying the sentence stimuli.

With this kind of sentence, where words contain jumbled letters, both competence levels are forced to focus at word level, performing syllable level chunking. MC and LC have the same number of chunks (18) at word group level and word level (represented by the black vertical border). However, when syllable chunking is taken into account, MC produces a total of 28 chunks, while LC produces 35 chunks. Although both MC and LC produce the same number of chunks at word group and word level, the syllable chunking that occurs within the words suggests that the processing of each may differ. The PLP values determined at the beginning of each word also demonstrate that each competence level engages different processes, which may be influenced by their language knowledge.

MC is most likely to refer to the Semantic System before copying the jumbled-letter word, in order to recall the actual spelling, hence the PLP value rising to 8. LC, being unfamiliar with the language, is not able to recognise the actual words and will decide to chunk at syllable level. Due to having the same number of chunks, both MC and LC produce the same median of PLP, with

values of 6. Although both competence levels produce the same PLP value, the differences lie in the different processes engaged during the copying process.

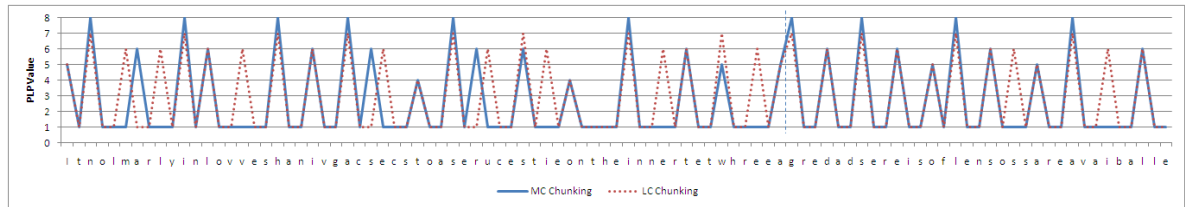


Figure 6.7 PLP values for a letters jumbled condition sentence

Figure 6.7 shows the PLP values for both MC and LC. The graph shows that the PLP value of MC has increased and at some points is larger than LC. MC's PLP value is significantly higher than for the previous jumbling conditions, indicating that this sentence type has required them to do more processing in addition to the syllable level processing. To conclude, it can be seen that MC has a comparable number of chunks to LC and that MC's PLP values have now increased. Thus, it is predicted that MC will have pause values similar to LC, although the pauses might refer to different processes, depending on the levels of knowledge.

#### 6.2.4.6 Predictions for Both Jumbled Condition (JWJL)

Finally, the sample sentence for the *both jumbled* type:

*"the it innertet nolmarly whree inlovves a hanivg gredad acsecs sereis to of a lensoss seruce are stie avaihalle on"*

This sentence is predicted to be the most difficult for MC to copy and completely meaningless to LC. MC would find it difficult to build up the sentence meaning, due to the jumbled words and distorted spellings. This might lead MC to copy without regard to meaning and might even cause them to chunk at syllable level, especially when word recognition is difficult.

The same phenomenon is predicted for LC. Due to their weak language knowledge, however, LC might decide to chunk at syllable level earlier than MC. Both MC and LC require greater processing time than for previous sentence types, resulting in longer overall pause lengths. This type of sentence might not be useful for assessing the cognitive processes in language processing, but is something that the present experiment will seek to elucidate further for the completeness of the experiment and tasks being explored.



Both Jumbled	t	h	e	i	t	i	n	n	e	r	t	e	t	n	o	l	m	a	r	i	y	w	h	r	e	e	i	n	l	o	v	v	e	s	a	h	a	n	i	v	g	g	r	e	d	a	d	a
MC Chunking	4					8					6				8		6					6					8		6				5		8		6			6						8		
Group of Words / Word / Syllable	4			1		8				6				8		6					6					8		6				5		8		6			6					8				
Letter	1	1		1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1			
Both Jumbled	t	h	e	i	t	i	n	n	e	r	t	e	t	n	o	l	m	a	r	i	y	w	h	r	e	e	i	n	l	o	v	v	e	s	a	h	a	n	i	v	g	g	r	e	d	a	d	a
LC Chunking	4					7				6				7		6			6			7		6			7		6		6		5		7		6		6		7		6		7			
Group of Words / Word / Syllable	4			1		7		6		6				7		6			6			7		6			7		6		6		5		7		6		6		7		6		7			
Letter	1	1		1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
Both Jumbled	c	s	e	c	s	s	e	r	e	i	s	t	o	o	f	a	l	e	n	s	o	s	s	s	e	r	u	c	e	a	r	e	s	t	i	e	a	v	a	i	b	a	i	l	e	o	n	
MC Chunking		6				8		6								4								6						5			5				8			6					5			
Group of Words / Word / Syllable		6				8		6								4								6					5			5				8			6					5				
Letter	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
Both Jumbled	c	s	e	c	s	s	e	r	e	i	s	t	o	o	f	a	l	e	n	s	o	s	s	s	e	r	u	c	e	a	r	e	s	t	i	e	a	v	a	i	b	a	i	l	e	o	n	
LC Chunking		6				7		6								4								7		6		6		5			7		6			7		6			6		5			
Group of Words / Word / Syllable		6				7		6								4								7		6		6		5			7		6			7		6			6		5			
Letter	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		

Figure 6.8 Sentence example for both jumbled condition

Figure 6.8 analyses the copying of the *both jumbled* condition and Figure 6.9 shows the overall outcome. There is a strong possibility that syllable level chunking will occur. As can be seen, both MC and LC have the same total number of chunks, 17. However, the processing of each chunk may differ, as seen from the PLP values at the beginning of each word. MC may try to comprehend the sentence and therefore require more processing, producing longer pauses. This is represented here by the PLP value of 8. LC, being unfamiliar with the language, is not able to recognise any words and is likely to decide to chunk at syllable level during the early stage of copying. Because of the many small chunks at syllable level, the medians of PLP for LC are the same as MC: a value of 6.

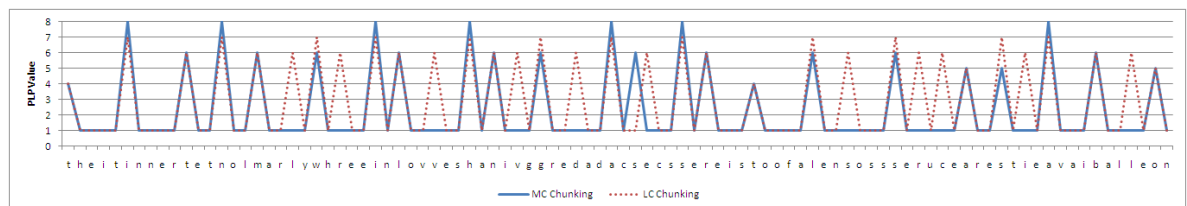


Figure 6.9 PLP values for a both jumbled condition sentence

Figure 6.9 shows the PLP values for both MC and LC. The results, however, are similar to those in Figure 6.7, where the chunking count and PLP values have increased for both MC and LC. *Both jumbled* is predicted to be the most difficult sentence condition, with a high occurrence of syllable chunking. It can be seen that MC and LC have large numbers of chunks and that MC's PLP values are higher or similar to LC, indicating the high demand of processing needed to perform the copying.

#### 6.2.4.7 Summary of All Predictions

To summarise the predictions made above, the median of PLP values and the number of chunks are compared. The results, which for the purposes of this analysis exclude the letter level processing, are presented in Table 6.3 below.

		Median of PLP		Number of Chunks at Word Group & Word Level Chunking		Number of Chunks at Word Group, Word & Syllable Level Chunking	
		MC	LC	MC	LC	MC	LC
Jumbling Type	No Jumbling	4	6	7	17	7	26
	Words Jumbled	4	6	8	16	8	25
	Letters Jumbled	6	6	18	18	28	35
	Both Jumbled	6	6	17	17	24	34

Table 6.3 Median of PLP results from processing the sample sentences

As shown, the *number of chunks at word group and word level chunking* is associated with the number of L2 pauses and the *number of chunks at word group, word and syllable level chunking* (i.e. *all chunks*) is associated with the overall distribution of pauses.

Overall, it can be concluded that the simulations divide the jumbling conditions into two groups: similar effects can be seen between *no jumbling* and *words jumbled* conditions, and between the *letters jumbled* and *both jumbled* conditions. From the analysis, *no jumbling* and *words jumbled* provide a greater differentiation of competence compared to *letters jumbled* and *both jumbled*. *It is different with respect to median of PLP, the number of chunks at word group, word level chunking and all chunks.* The inclusion of syllable level chunking in the chunk counting better differentiates MC and LC.

Therefore, the results of simulations suggest that *no jumbling* and *words jumbled* may be better stimuli than the other two, as they are able to distinguish the levels of nature of language processing.

## 6.3 Method

### 6.3.1 Participants

Twenty native Spanish speakers were recruited, with an age range of 20 to 35 years old. None had language impairments or handwriting problems. All the participants had English as their second language and had graduated from high school. The participants came from different Spanish-speaking countries: seven from Mexico and thirteen from Spain. Six of them were primary school teachers who had come to the University of Sussex for an English course, while the rest were postgraduate students from the same university. In both countries, English is taught at school from a primary level, but in Mexico those who wish to improve their English can also attend a special tuition centre. There is a considerable difference in English competence levels between those who attended the special tuition course and those who did not.

### 6.3.2 Apparatus

The same methodology and apparatus as Experiment 1, as explained in Section 5.2.2, is applied in this experiment.

### 6.3.3 Procedure

The procedure followed here is the same as in the previous experiments. The participants were asked to copy the sentence stimuli as quickly and as accurately as possible. The stimuli were visible at all times, but the participants were not allowed to read them in advance. Participants were only allowed to read the stimuli and begin copying when the researcher said 'start'. The participants were trained to begin each sentence with a 'hash' (#) to ensure that the writing process is well underway before the first stimulus letter was written.

The participants were asked to copy 20 stimuli each, comprising 12 sentence stimuli and 8 name writing (NW) tasks. The order in which they copied these stimuli is explained in Figure 6.10.

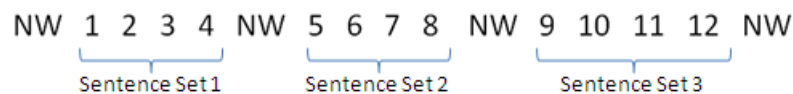


Figure 6.10 The order of stimuli

The 12 sentences were divided into three sets of four, wherein all four jumbling types were represented. In between each set is a name writing (NW) task, each consisting of upper case name writing and lower case name writing, which are used as baselines for each participant. The NW baseline was found in previous experiments to give the highest correlation of the tested baselines, so this experiment focuses only on NW. The arrangement is done in such a way so as to sample NW throughout the trials.

Of the 12 sentence stimuli, each sentence has 4 different types of jumbling, totalling 48 different stimuli. Each participant will be required to copy 12 different sentences over three sentence sets, each set containing a sentence from each jumbling type. This is to make sure that the sentences are counter-balanced and the jumbling type is distributed equally across all participants and sentence presentation order. This is illustrated further in Figure 6.11.

		Participant Number																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
		Grouped by the first trial received																			
		G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4
		UPPER CASE NAME WRITING																			
SenNo		upper case name writing																			
Sentence Set	1	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d
	2	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a
	3	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b
	4	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c
		UPPER CASE NAME WRITING																			
		upper case name writing																			
Sentence Set	5	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d
	6	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a
	7	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b
	8	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c
		UPPER CASE NAME WRITING																			
		upper case name writing																			
Sentence Set	9	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d
	10	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a
	11	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b
	12	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c
		UPPER CASE NAME WRITING																			
		upper case name writing																			
LEGEND:																					
SenNo - Sentence Number																					
a	No Jumbling - NWNL	G1	First trial is no jumbling																		
b	Words Jumbled - JWNL	G2	First trial is words jumbled																		
c	Letters Jumbled - NWJL	G3	First trial is letters jumbled																		
d	Both Jumbled - JWJL	G4	First trial is both jumbled																		

Figure 6.11 The design of the stimuli order

Figure 6.11 demonstrates the order of the stimuli given to all 20 participants. The *jumbling type* is rotated for each participant, but the *sentence presentation order* is maintained in the same

position. For the ease of understanding the design of the stimuli order, particularly for this section, a label was assigned for each jumbling condition, as explained in the legend. These jumbling types are rotated across all the sentences for each participant in order to give a balanced distribution of the stimuli type across all participants. To get an even distribution of sentence number (SenNo), the sentence set, which contains the four jumbling conditions, was rotated. The *first trial* for an individual is defined as the first sentence stimuli received in the sentence set. As shown, the rotation starts at every 1<sup>st</sup>, 5<sup>th</sup> and 9<sup>th</sup> sentence, the beginning of every sentence set.

		Grouped according to the first jumbling type received			
		G1	G2	G3	G4
Jumbling Type	No Jumbling (a)	1 <sup>st</sup>	4	3	2
	Words Jumbled (b)	2	1 <sup>st</sup>	4	3
	Letters Jumbled (c.)	3	2	1 <sup>st</sup>	4
	Both Jumbled (d)	4	3	2	1 <sup>st</sup>

Table 6.4 Trial order number for each participant

For each rotation, participants were grouped according to the first jumbling type received. Table 6.4 summarises the four groups of participants with the jumbling type order represented by the number; *trial order number*. Participants who received *no jumbling* (a) in their first set of sentences (i.e. 1<sup>st</sup>) are put in G1 (Group 1). G2 participants received *words jumbled* (b) as their first set, G3 participants received *letters jumbled* (c) and G4 *both jumbled* (d). There were five participants for each group.

Overall, the 20 participants produced 240 sentences and 160 NW sets of data.

Upon completing the sentence-copying tasks, the participants were asked to take two independent tests of language proficiency, as a means to measure their English competence.

#### 6.3.4 Language Performance Test

This experiment extends the number of language components tested by introducing a vocabulary test in addition to the ODT grammar test. This section explains the tests used.

### 6.3.4.1 The English Performance Test (ODT)

As in Experiment 1, this experiment used the ODT Advanced test as one of the language proficiency tests, focusing on grammar knowledge. The motivation behind using ODT is described in Section 5.2.5 (see Chapter 5).

### 6.3.4.2 The Vocabulary Size Test (VST)

There are a number of established vocabulary test tools (e.g. Nation & Beglar, 2007; Laufer, 2005) that have been used in research for the purpose of validation (Chapter 2). Taking into account a variety of factors, including the test availability, and the difficulty and suitability of the question types, the Vocabulary Size Test (VST) by Nation (2007) was adopted.

The VST tests word families of the English language based on the British National Corpus (BNC) 10 million word list. It contains 140 test items, 10 from each 1000 word level, and each item is given a score of 100. This gives a maximum score of 14000. The word families were chosen with care based on the *Level 6 criteria* of Banner and Nation's (1993) scale of levels. These criteria include, among others, regularity, frequency, productivity and predictability. The goal for a learner who aims to understand non-simplified spoken or written texts is to reach at least the 8000 word family. Scoring below 8000 implies lower language proficiency.

### 6.3.5 The Overall Structure of this Experiment: The Variables

Figure 6.12 shows the overall structure of this experiment.

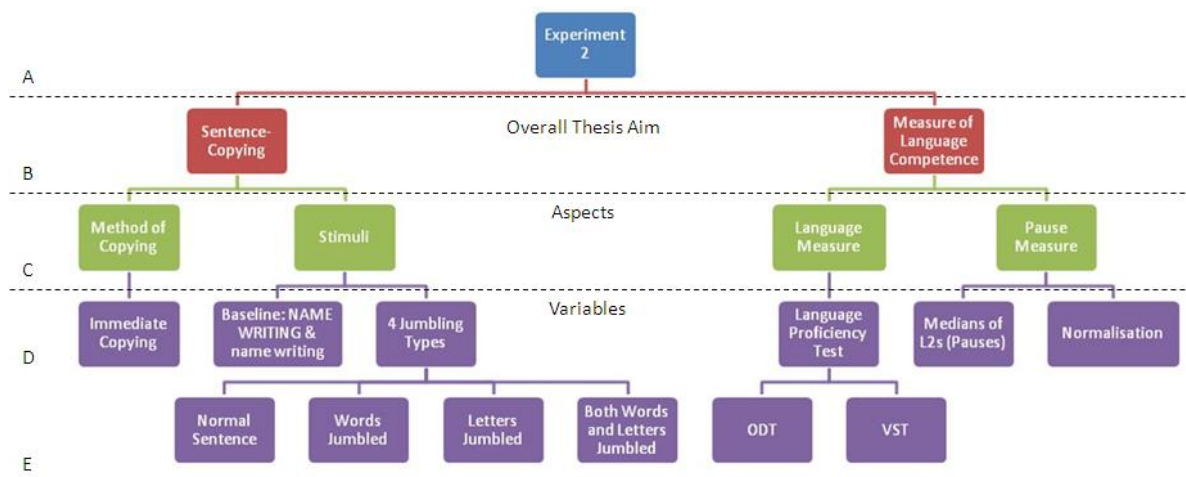


Figure 6.12 The overall structure of this experiment

Four aspects presented at level C group a number of components (level D and E) that are important to this study. Some of these variables were established in previous experiments, and some were developed in this chapter. Variables used in previous experiments are (1) Immediate Copying (IC), (2) the ODT, (3) median of L2 pauses and (4) normalisation. The additional variables introduced in this chapter are (1) upper and lower case name writing baseline (refined from previous experiments), (2) sentence stimuli produced based on the four jumbling types, and (3) a vocabulary test to validate word recognition, the VST.

### 6.3.6 Analysis

All the raw data were extracted using TRACE in order to identify the pauses between each stroke. Instead of mapping each pause value to the letters of the sentences and then determining the pause levels manually, (as in Experiment 1), PLET (van Genuchten, 2009) was used to locate the pauses and perform the L2 pauses median calculation. The pauses were then correlated against the ODT and VST scores. The analyses were conducted in a top-down fashion, from the greatest aggregation down to specific cases:

1. Effects of overall pause values irrespective of jumbling type.
2. Effects of overall pause values aggregated for each jumbling type.
3. Effects of jumbling type and order of presentation aggregated for just the first trial received.
4. Effects focused on the first sentence set (of three sentence sets), for reasons to be explained below.

For the purpose of improving the measure, the use of normalisation was tested. The result of normalising is compared to the results that focused on the first sentence set (previous analysis).

## 6.4 Results

The compilations of scores from the language proficiency tests were correlated to observe any relationship between the two tests. A *Pearson correlation coefficient* ( $r$ ) was used to measure the strength of the relationships between the two tests, ODT and VST, and between the tests and the L2 median pauses (from sentence-copying). In simple terms, participants who achieved good scores in the ODT are predicted to achieve similar good scores in the VST, but as the tests are

measuring two different components, they wouldn't produce a strong correlation. In terms of the sentence-copying, it is predicted that having a great familiarity with the language produces good test scores i.e. many short pause lengths, but a low familiarity level would produce poor test scores i.e. many longer pause lengths.

#### 6.4.1 Validating the Independent Language Tests: ODT & VST

Figure 6.13 shows the correlation between ODT and VST, in order to confirm that both tests are measuring different components of language.

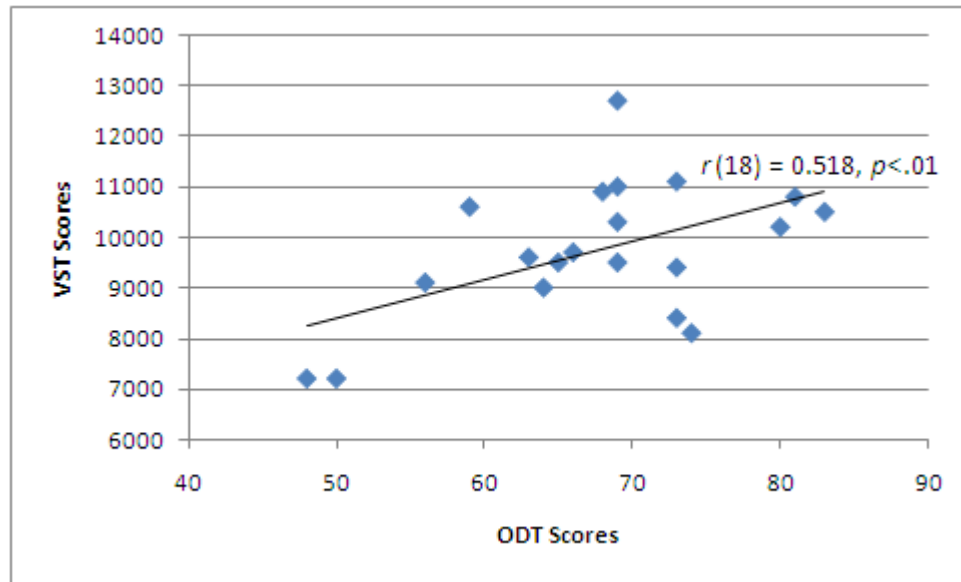


Figure 6.13 The correlation between ODT and VST

As can be seen from Figure 6.13, there is a good range in the competence level scoring from the 20 participants in this experiment, between 48 and 83 for ODT and between 7200 and 12700 for VST. The correlation between the vocabulary and the grammar test for these 20 participants is  $r=0.518$  ( $p<.01$ ). This is reassuring, in that they both show a roughly similar rank of participants – with MC achieving good scores for both and LC to low scores for both – but at the same time, the two tests are shown to be addressing different language components.

#### 6.4.2 The General Effects of Copying Different Jumbling Conditions

Table 6.5 presents the correlation results of the (1) *Overall Pause Values* and (2) *Jumbling Type*. The *Overall Pause Values* shows the overall pauses correlated against ODT and VST scores. The *Jumbling Type* shows the correlation results between the mean values of the L2 median pause



values of the jumbling type against the test scores. The Overall Pause Values comprise data from 20 participants, as does each jumbling condition.

Jumbling Type		ODT	VST
	Overall Pause Values	0.100	0.314
	No jumbling (a)	0.170	0.046
	Words jumbled (b)	0.278	0.067
	Letters jumbled (c)	0.242	-0.041
	Both jumbled (d)	0.433*	0.079

Table 6.5 Correlations between the English tests and pauses (\* $p < .05$ )

*Overall Pause Values* shows weak correlations for both ODT and VST. When the data are aggregated according to the jumbling condition, the correlation results are still weak, although the values for the jumbling condition with ODT has increased compared to the *overall pause values*. It is interesting that when the pauses are segregated according to the jumbling condition, *both jumbled* was a significant correlation. On the other hand, the correlation values for each jumbling type against the VST shows a much weaker value compared to the *overall pause values* for VST. The correlation for jumbled letters is slightly negative.

#### 6.4.3 Data Aggregated According to the Jumbling Conditions & First Jumbling Type Received

Figure 6.14 presents the aggregated data according to the first trial jumbling condition, producing four groups with 5 participants each. The data take the overall pauses from all sentence sets, irrespective of the specific sentence set. Each data point represents a correlation between the pauses of the particular jumbling type and the tests. A *sentence set*, as explained in Section 6.3.3, includes four sentences – a different sentence for each of the four different jumbling conditions (*no jumbling*, *words jumbled*, *letters jumbled* and *both jumbled*) separated by the NW tasks.

The naming convention consists of [Group] [group number] [(jumbling condition they received for first trial)], for example G1(a). Figure 6.12 and Table 6.4 in Section 6.3.3 provide a reference to the arrangement of the jumbling condition. To demonstrate, Figure 6.14 provides data aggregated by the *first jumbling type* received by participants, segregated by the two tests. Each *first jumbling type* group consists of 5 participants who performed the four different jumbling conditions. Each group differs by the first trial jumbling type. For example, G1 (*no jumbling*), has 4 data points, one for each jumbling condition, performed by the 5 participants in this group, and the first trial starts

with *no jumbling* (blue solid line) followed by *words jumbled* (red dotted line), *letters jumbled* (green small dashed line) and *both jumbled* (purple larger dashed line). G2 (*words jumbled*) then provides 4 data points from another 5 participants, and the first trial is *words jumbled* (red dotted line), then *letters jumbled* (green small dashed line), *both jumbled* (purple larger dashed line) and *no jumbling* (blue solid line). The same pattern applies to G3(c) and G4(d).

An illustration of how one jumbling type is positioned across the diagram is as follows: in G1 (*no jumbling*), participants performed *no jumbling* condition as the first trial. In G2 (*words jumbled*), participants performed *no jumbling* condition as the fourth trial, after the completion of *words*-, *letters*-, and *both jumbled*. In G3 (*letters jumbled*), participants performed *no jumbling* condition as the third trial, after the completion of *letters*- and *both jumbled*. In G4 (*both jumbled*), participants performed *no jumbling* condition as the second trial, after the completion of *both jumbled*.

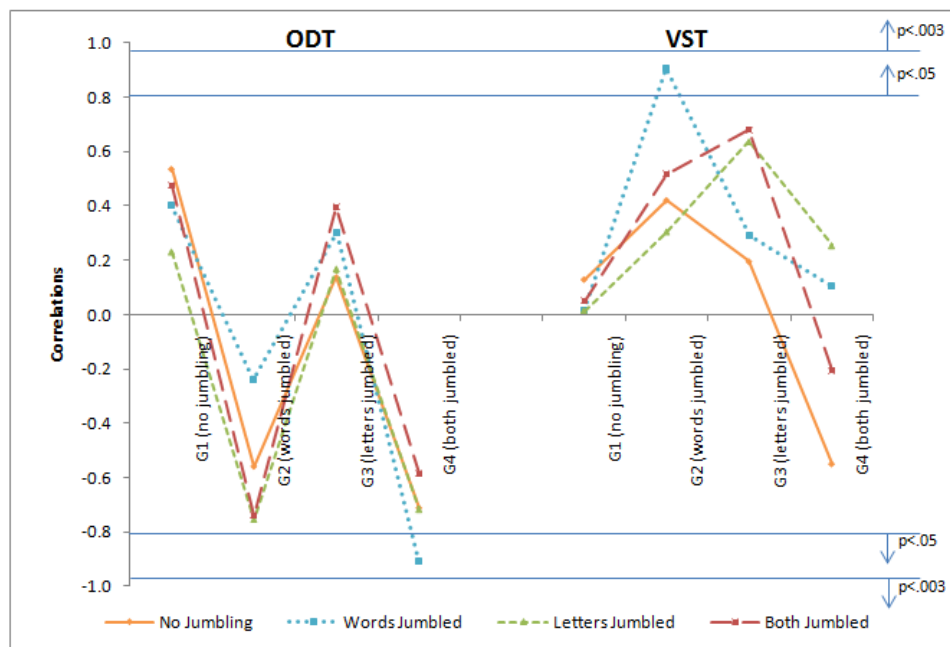


Figure 6.14 Overall results (all sentence sets) of copying the different jumbling conditions, aggregated by the first jumbling type received

It is apparent from this graph that the combination of pauses from all three sentence sets reveals few strong correlations, although interestingly the correlation values are greater than in Table 6.5 in many cases. G4 (*both jumbled*) which had a strong positive correlation previously (c.f. Table 6.6), has now, surprisingly, shown a strong negative correlation. Although the overall data point

correlations have improved, the Bonferroni procedure used to control for Type I error across the sixteen correlations comparisons ( $\alpha' = .05/16 = .0031$ ) for each test, was not able to provide any significant correlations.

The improved results suggest that breaking down the aggregation further may benefit the analysis. This will be carried out in the next section

#### 6.4.4 Data Aggregated by Sentence Set, Focusing Only on Sentence Set 1

Figure 6.15 focuses only on the first sentence set and excludes the second and third sentence sets from the analysis. The results for some groups and jumbling conditions now have some stronger correlations. The results of second and third sentence sets (not presented here) failed to show strong correlations, which strengthens the view that the first sentence in each set is particularly meaningful.

The data for G1 (*no jumbling*) for both ODT and VST are interesting because they suggest an effect caused from the order of jumbling type presentation, which has caused the correlation values to improve. The correlations have now widened compared to the previous analysis, where the values were more clustered. G2 (*words jumbled*) for ODT has also improved in its correlation values. Other correlation values have either maintained their position or have changed only slightly. Although the overall data point correlations have improved, the Bonferroni procedure used to control for Type I error across the sixteen correlations comparisons ( $\alpha' = .05/16 = .0031$ ) for each test, was still not able to provide any significant correlations.

In order to compensate for any individual differences, data normalising will now be examined.

#### 6.4.5 Normalisation

The normalisation was carried out using *pause difference* from Experiment 1. Pause difference involves deducting the median of L1 (letter) pauses from the median of L2 (word) pauses for each individual and was shown to be the best method. Figure 6.16 presents the normalised data, using the L1 median pauses values obtained from the sentence stimuli. The normalisation was also conducted using the median of L1s from the NW baseline tasks and similar results were found.

The analysis has shown no clear pattern of improvement and correlations. There are substantial changes in some correlations, but not others. Thus, there is no evidence to say whether normalisation has worked, which implies that the process of subtracting the medians of L1s is causing an effect additional to that of correcting for individual writing speeds. However, the Bonferroni procedure used to control for Type I error across the sixteen correlations comparisons ( $\alpha' = .05/16 = .0031$ ) for each test, was only able to provide one significant correlation.

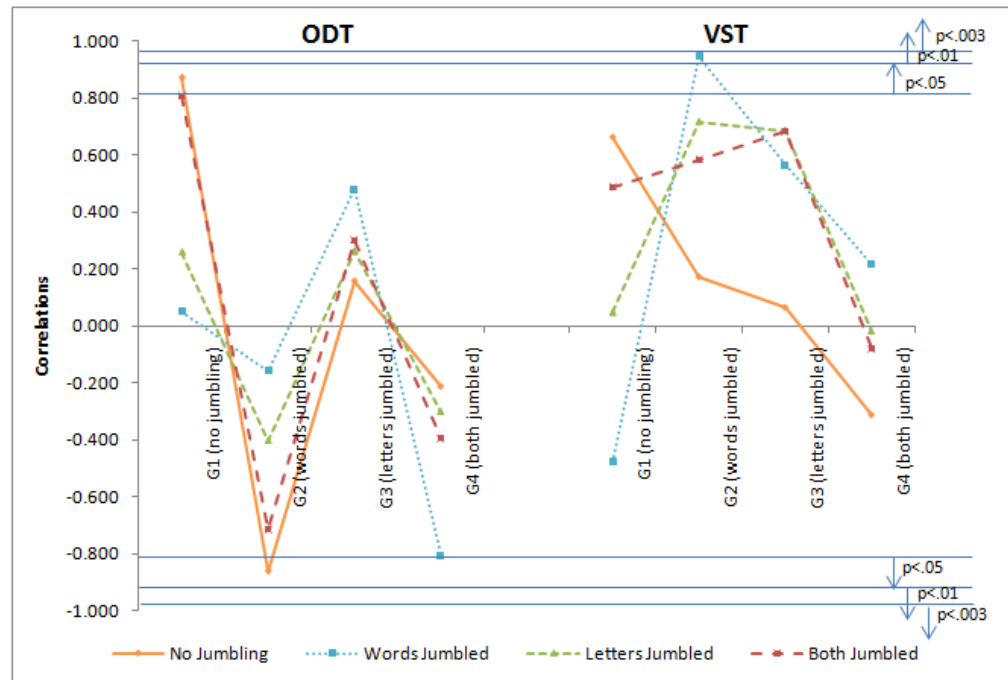


Figure 6.15 The results from copying the first sentence set, aggregated by the first jumbling type received

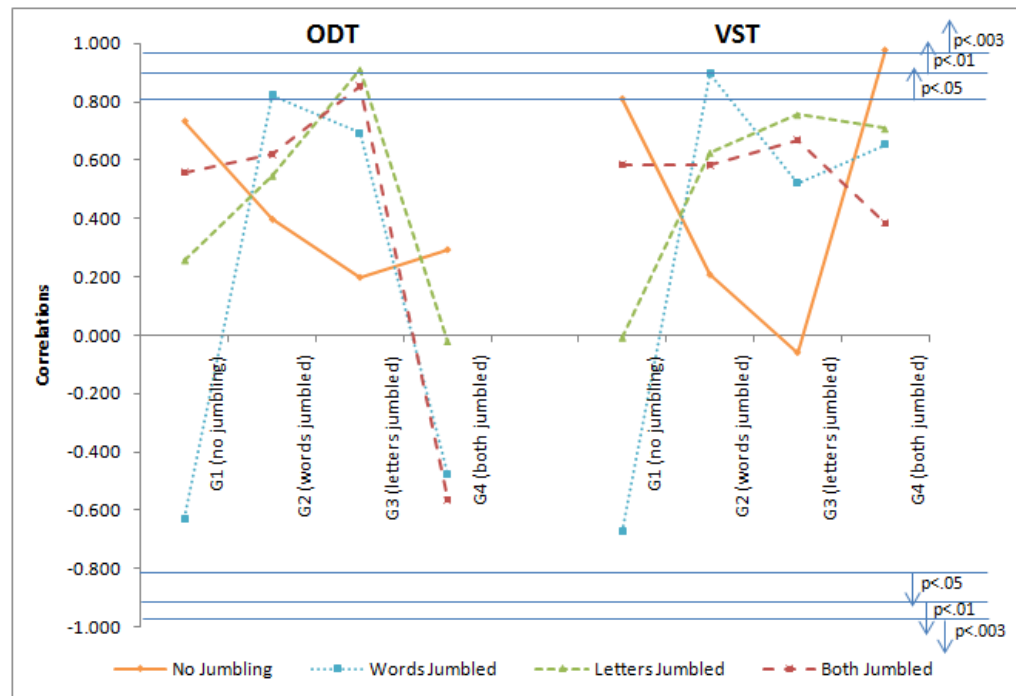


Figure 6.16 The normalised results of the first sentence set, aggregated by the first jumbling type received

Overall, it is shown that the range of correlation values has widened across the stages of analysis, increasing gradually over Figures 6.14, 6.15 and 6.16; especially for G1 (*no jumbling*) where we can see the difference between the first trial (*no jumbling*) and the second trial (*words jumbled*), for both ODT and VST. The overall correlation pattern is quite similar across the two tests and the groups of participants (according to the first trial jumbling type received). The correlations are more clustered in Figure 6.14 for the whole of ODT, but only in G1(*no jumbling*) for VST.

## 6.5 Discussion

The aim of this experiment was to investigate the effect of jumbled sentence stimuli for assessing the cognitive processes of language processing. More specifically, this study investigated four jumbling conditions, namely: *no jumbling*, *words jumbled*, *letters jumbled* and *both jumbled*. It was predicted that each of these jumbling conditions would affect the copying differently. Furthermore, it was predicted that these jumbling conditions would be able to assess different language components: grammar knowledge and word knowledge.

Overall, the findings suggest that one effect of jumbling order presentation is that the first trial of a particular jumbling type received determines the strategy of copying applied to the subsequent jumbling types. In other words, the first instance of a particular jumbling type affects how the participant processes subsequent sentences with a strategy suited to that jumbling condition applied. Thus, performing the copying is affecting the levels of processing, (i.e. word group-, word-, syllable- and letter-level chunking).

For example, if a low competence person receives *letters jumbled* as the first trial, the chunking size would most probably be at word and syllable level. This increases the likelihood that the same chunking process will occur with the next jumbling condition received: *both jumbled*. This might not make much difference as applying the same strategy would be suitable for *both jumbled*. However, the case is more obvious when the first trial is *both jumbled*, followed by a *no jumbling* condition. The use of word and syllable level chunking for *both jumbled* might not be suitable for the *no jumbling* condition, especially when the grammatical structure and spellings are in their correct forms. However, the strategy from copying the *both jumbled* has been applied to copying the *no jumbling* condition. This may be the reason for the inconsistent patterns of correlation and weak correlations for some data points.

This finding is important because it explains the reasons for the objective of this experiment, to find suitable sentence stimuli for sentence-copying, not being achieved. Further, it has highlighted the need to control the nature of the copying tasks and the activity of copying. This extends the overall aim of the study to develop a novel testing method, by suggesting that there is a need to carefully design tasks in order to avoid such effects.

The findings did, however, find some effects of the different jumbling conditions, as shown by a number of strong correlations. With these in mind, there is certainly a need for more study to refine the method further. Despite the limitations of the experiment, detailed analysis produced a number of useful findings which contribute to the overall aims of this study. These are explained in the subsections that follow, seeking to answer (1) if ODT and VST are measuring different language components; (2) whether jumbling is a good way to design stimuli; (3) whether normalisation could improve the correlation values, and (4) why certain jumbling conditions did not work.

### 6.5.1 Are ODT and VST Measuring Different Components?

The results of this study show that the ODT and VST tests measure different components and this is supported by the weak correlations produced between the two tests. ODT and VST were used as an independent measure of language proficiency for validating the suitability of the sentence stimuli. The aim of ODT is to assess grammar knowledge and of VST to assess vocabulary size.

This experiment found different correlation patterns produced between each test (ODT and VST) and the L2 median pauses from the copying of each jumbling condition indicating different effects from the sentences or the way one copies the sentences. This experiment, which focuses on the four jumbling conditions, provides the potential of measuring grammar, when the structure is manipulated, and of measuring word knowledge, when the letters in the word are transposed.

These findings support the idea that they are measuring different language components and these are important to confirm that the applications of the two tests are meaningful to this study and that they do not assess the same components. In conclusion, the application of ODT and VST is shown to be an appropriate choice for measuring the two language components; grammar knowledge and vocabulary knowledge.

### 6.5.2 Is Jumbling a Good Way to Design Stimuli?

This experiment suggests that jumbling may be a good way to design stimuli, based on the strong correlations found with certain jumbling conditions. With the relevant kind of manipulation employed, the jumbling conditions were able to manipulate the grammar structure and word knowledge, and thus assess, the grammar and vocabulary level.

The predictions made through the simulations of MoC suggested that the *no jumbling* and *words jumbled* conditions may provide suitable sentence stimuli, as they were able to differentiate the levels of competence, MC and LC. *Letters jumbled* and *both jumbled*, on the other hand, show no difference between the two levels of competence. However, this prediction could not be compared with the actual data because of the issues mentioned earlier. Furthermore, an attempt was made to circumvent the problem by considering just the first sentence of each jumbling type but the number of participants in each group does not provide a good range of language performance scores and so was too small to give reliable data.

In comparison to the findings of Experiment 1, where the correlations achieved were all negative, Experiment 2 produced a mix of positive and negative correlations. This suggests that there might be something interesting occurring such as the kind of sentences used in the experiments and the different group of participants (Malaysian and Spanish). Again, the limitation of the study prevents further analysis.

Although the results were not able to provide strong correlations for all jumbling conditions because of the limitation found, this does not mean that the exploration fails. This limitation suggests a need for further investigation of the jumbling conditions, but with an improved sentence stimuli design. It also suggests that there is a need to control the copying of the stimuli in order to avoid the possibility of participants employing the same strategy used in the previous sentence onto the next sentence. The next chapter will therefore look into the exploration of task conditions that are aimed at controlling the copying activity.

Because of the limitations encountered, the copying of the first and second trial of each jumbling type was examined. This examination was carried out in order to investigate and understand the processes occurring during the activity of copying. It is predicted that there is a zero possibility for sentence order effect on the first trial, as this would be the first occasion that the participants deal with the jumbling sentences. The second trial however, is observed to see if the copying in the first trial is affecting the copying of the second sentence stimuli. These explanations are discussed in more detail in Section 6.5.5.

Despite the issue of the effect of jumbling order presentation, there are still a number of strong correlations in certain cases, which indicate the potential of using the jumbling condition. For example, in the case of G4 VST (refer to Figure 6.17), doing sentences with *both jumbled* condition first does not affect sentences with *no jumbling* condition which came next. This may suggest a good strategy that can be practised because having this kind of arrangement (i.e. *both jumbled* followed by *no jumbling*) is able to separate the strategy of copying.

### 6.5.3 Did the Normalisation Appear to Improve Measures?

This experiment did not detect any evidence of a clear pattern of improvement after applying *normalisation*. There are a number of substantial changes in correlation, but only for some data



points. The technique used in Experiment 1 was found to be useful in distinguishing the effective sentence type.

Although corrections to the data have been done, the results are still not strong, which indicates that there might be other processes going on, which hamper the pauses. For example, after performing normalisation for the data points of G2(b)<sub>ODT</sub>, the correlation values swapped from being negative correlations to being positive correlations. This phenomenon could not be explained because of the limitations found, which may have affected these occurrences. However, these differences are noted for the purpose of understanding the effects of each jumbling condition. Thus, it cannot at this point be established whether normalisation is a useful approach. On the contrary, performing the normalising has highlighted that an individual's speed of copying may vary in two ways from their standard writing speed: (1) it may speed up (fast-copying), because of a greater level of familiarity, or (2) it may slow down (slow-copying), because of the difficulty of the sentence stimulus. 1 can be seen from the pause distribution that there are many short pauses and that they are almost a flat line, whilst, 2 can be seen from the many long pauses that can be clearly distinguished compared to familiar letter copying.

Although the overall impact of normalisation may not be great, due to the low number of significant correlations, the improved first trial data for some of the jumbling conditions may be able to help in understanding the processes taking place during the copying of the different jumbling types. It also helps the understanding of the effects that occur because of the ordering of jumbling type, which determines the strategy of copying the rest of the stimuli. Making predictions is therefore explored, to understand what processes are occurring during the copying, as explained in detail in Section 6.5.5.

#### 6.5.4 Why Did Both Words Jumbled and Letters Jumbled Not Work?

Contrary to expectations, this study did not find a consistent significant correlation for both *words jumbled* and *letters jumbled*. This was probably affected by the order effect of which sentence type is seen first, which determines the copying of the subsequent sentences. Moreover, it is not possible to make comparisons given the limited number of participants in a group with a disproportionate range of language performance scores. This may be one of the reasons the correlations were not strong. The group of participants, i.e. Spanish speakers, recruited for this experiment also differs from that in the previous experiments, Malaysians. This might affect the

interpretation of data as both language speakers might have different strategy in processing English language which may be influenced by their own culture .

These findings, need to be explored further, as discussed in previous sections. The ability to provide strong correlation between the first trial and the MoC predictions indicates a possibility that this jumbling condition might work. Thus, the condition will continue to be investigated in the next experiment.

Because of the same limitations that have impeded the analyses and results across all the discussions of this experiment, the decision was taken to attempt to understand the processes that occur during the copying of the first and second trial, which will be explained in the next section.

### 6.5.5 What is Happening During Copying in the First & Second Trials?

The first trial for each jumbling condition will not be affected by the order of the jumbling type, however, the first sentence trial of the first sentence set may, however, impact on the copying strategies on the rest of the sentences across all the sentence set. Because of the effects of copying strategies, the first two trials that participants first received is observed in order to analyse the processes that might take place which affect the decision of copying strategy for the subsequent jumbling type. Therefore, the sections that follow will explain the impact of each jumbling type when used as the first trial. However, the descriptions here only considers cases where the ODT and VST have a large range of scores, hence does not include G4(d).

Participants adopt their initial copying strategy based on the type of sentence received in their first trial. This strategy will then be applied to the first and second trials. As this is likely to change for the third and subsequent trials, it is only safe to make inferences about the first two trials for a particular group. The assumptions discussed in the sections that follow are graphed in Figure 6.16.

#### 6.5.5.1 Explanations for G1(a):

This section interprets G1(a), where the participants' first trial in copying is a *no jumbling* sentence (a) followed by a *words jumbled* sentence (b). An attempt is made here to explain the reasons behind the sudden change in Figure 6.17, from  $r_{VST}(3)=+0.811$  ( $p<.05$ ) and  $r_{ODT}(3)=+0.734$  correlation values (diamond marker) to  $r_{VST}(3)=-0.672$  and  $r_{ODT}(3)=-0.629$  values (square marker).

Comparing VST and ODT, the changes in its correlation values is apparent between the first trial and the second trial. Indeed it is interesting that the positive correlations in the first trial have changed to negative correlations in the second trial. This effect is maintained across the two trials.

Participants in group G1(a) (first trial=*no jumbling*) initially (in the first trial) attempt to code the sentences for meaning.

Competent participants may process meanings more deeply than the less competent. The former may spend more time than the latter in processing words because they need to:

(a) access the meaning of the words

(b) relate the words grammatically to each other.

Hence, their L2 pauses may be longer, both with respect to (a) VST and (b) ODT.

Participants in group G1(a) (first trial=*no jumbling*) again attempt to code the sentences for meaning on their second trial, with sentence type 'b' (*words jumbled*).

The more extensive vocabulary and understanding of competent participants may help them realise that the sentence structure is grammatically incorrect; therefore, they will process the sentence at word level which will generally be faster than LC.

Hence, they will need less time than the less competent.

#### 6.5.5.2 *Explanations for G2(b):*

This section interprets G2(b) where the participants' first trial in copying is *words jumbled* sentence (b) followed by *letters jumbled* sentence (c). In this condition, the values slightly decrease from the first trial (square marker in Figure 6.17) to the second trial (triangle marker): from  $r_{VST}(3)=+0.897$  ( $p<.05$ ) and  $r_{ODT}(3)=+0.823$  ( $p<.05$ ) correlation values to  $r_{VST}(3)=+0.626$  and  $r_{ODT}(3)=+0.547$  values. This case differs from G1a. Here, an attempt is made to explain what might affect the decrease in correlation value. Both VST and ODT show some similar effects.

Participants in group G2(b) (first trial=*words jumbled*) initially (in their first trial) attempt to code the sentences for meaning.

Competent participants may process meanings more deeply than the less competent.

The former may spend more time than the latter processing words because they need to:

- (a) access the meanings of the words
- (b) relate the words grammatically to each other.

Hence, their L2 pauses will be longer, with respect to both (a) VST and (b) OST.

Participants in group G2(b) (first trial= *words jumbled*) again attempt to code the words for meaning on the second trial, with sentence type 'c' (*letters jumbled*).

The better vocabulary and understanding of competent participants will help them decode the words faster, in terms of reading and recognising the jumbled letters.

Competent participants, however, will find it difficult to copy the words in big chunks because of the orthographic condition of the jumbled letters.

They may spend more time than the less competent participants processing the words because they need to:

- (a) control the automaticity through which they tend to write words with the correct spelling, instead copying the letters exactly as presented
- (b) be extra careful in copying, and hence might adopt a strategy of copying at syllable level, as they are familiar with syllables

Overall, they will need more time than the less competent.

#### 6.5.5.3 *Explanations for G3(c):*

This section interprets G3(c) where the participants' first trial in copying is *letters jumbled sentence* (c) followed by *both words and letters jumbled sentence*(d). In this condition, the values slightly decrease from the first trial (triangle marker in Figure 6.17) to the second trial (cross marker), from  $r_{VST}(3)=+0.756$  and  $r_{ODT}(3)=+0.911$  ( $p<.05$ ) correlation values to  $r_{VST}(3)=+0.670$  and  $r_{ODT}(3)=+0.855$  ( $p<.05$ ) values. Here, an attempt is made to explain what might affect this slight decrease of the correlation value from the first trial to the second.

Participants in group G3(c) (first trial= *letters jumbled*) initially (in their first trial) attempt to code the sentences for word meaning or spelling.

Competent participants may process meanings more deeply than the less competent.

The former may spend more time than the latter in processing words because they need to:

- (a) be very careful in copying the words with different orthographic features, even though they can read them as they would normally and are able to recognise them automatically
- (b) chunk the words into smaller sizes in order to make copying easier and more manageable
- (c) relate the words grammatically to each other.

Hence, their L2 pauses will be longer, with respect to both (a) VST and (b) ODT.

Participants in group G3(c) (first trial=*letters jumbled*) on their second trial, with sentence type 'd' (*both words and letters jumbled*), attempt to code the words in a manner similar to the first trial. Not much difference can be observed, however, because of the similar condition of the jumbled letters.

Competent participants may spend more time than the less competent ones in processing words because they need to:

- (a) be very careful in copying the words with different orthographic features, even though they can read them as they would normally and are able to recognise the words automatically
- (b) chunk the words into smaller sizes in order to make copying easier and more manageable

At this stage, however, competent participants might have given up trying to relate the words grammatically.

Hence, their L2 pauses will be longer, with respect to both (a) VST and (b) OST.

#### 6.5.5.4 Conclusion of the Process Explanations

Based on the assumptions made above, it can be concluded that the first and second sentence trials might affect the copying performance of an individual for the rest of the sentences. The type of jumbling condition that they encounter first affects how they process the subsequent sentences. Because of the 'immediate copying' method, strategies begin to be developed during the process of copying. It is not surprising that the first sentence set produces a more revealing data set than the second and third, because participants have no idea on the kind of sentences they will be asked to copy. This can be clearly seen on the first and second trial of the first sentence set, as explained above.

#### 6.5.6 Limitation of the Experiment

The order in which the sentence jumbling conditions are given to the participants limits the study. Although an attempt has been made to produce a counterbalanced stimuli order, this is not an easy process. The study has tried to control the stimuli design, but this has restricted the randomisation of the sentence.

The limitation can be explained as: *no jumbling* is always followed by *words jumbled*, *letters jumbled* and *both jumbled* conditions, in sequence and then rotated, returning to *no jumbling*. Although the first jumbling condition varies, the order is consistent. This restriction has contributed to the findings of the limitation; it may not have been discovered if the sentence order presentation has been counterbalanced from the very beginning. This is indeed a valuable finding which assists in a tighter sentence stimuli design for the next experiment.

#### 6.5.7 Implications of the MoC

As an implication to the model, the limitation found shows that the first jumbling type seen in the sequence of sentences with different jumbling conditions affects the participant's decision in the choices of what copying method to adopt for the rest of the sentences. Observations made show that when someone is given a normal sentence to copy, they adopt a similar strategy of copying a normal sentence onto copying jumbled words, jumbled letters and both jumbled. This of course is a bit of a struggle, especially when words are not presented with normal spelling (as seen in the first sentence), hence the effects are shown on the pauses. The same occurs for other conditions, according to which jumbling condition is seen first. For example, when someone sees the

conditions of both jumbled first, the copying method adopted for the next type of sentence (in this case, the normal sentence) is likely to be the same method as used for both jumbled. This would eventually affect the overall pause distribution. Hence, the initial aim to see the effects of different jumbling conditions individually could not be seen.

Therefore, what actually occurs during the process of copying when the first sentence is a normal sentence, (followed by words jumbled, letters jumbled and both jumbled) a competent person may carryout chunking at word level or group of words level for that first sentence. When given the words jumbled condition (as the second sentence), this person might struggle in building up the meaning of the whole sentence. Even though they are still able to group words together, the effort to make up the whole meaning (which costs some extra milliseconds) may be unsuccessful. Those who are less familiar with the words may chunk the normal sentence at group of letters or syllable level, so when given the words jumbled condition they will adopt the same chunking process, and not refer to semantic processing. In terms of the MoC, this means that the copying process can be quicker or slower than the normal sentence, depending on the need to refer to semantic processing.

Given that the subsequent sentence is made of jumbled letters, the effort to group words would be more difficult, forcing the copier to focus on each word individually. The chunking level has now been downgraded to individual words, because of the jumbling conditions. In the MoC, because of the different presentation of the word, the individual might be forced to refer to semantic processing in order to search for some recognition of that particular word, to make copying easier. This itself has already added a few milliseconds to the normal copying processes, hence the copying process may be longer than for the normal sentence. A similar process would occur with other sentence conditions.

As a conclusion, given the fact that the results provide a limitation for observing the effects of jumbling on the copying processes, applying the MoC confirms that the actual processing of copying is influenced by the first jumbling type seen.

## 6.6 Conclusion

In this experiment, four jumbling conditions have been considered in order to establish a suitable sentence type. Although the effects of different jumbling types at different positions have been identified through observed effects, especially the effects on the first and second trial, the present data have not been able to distinguish a suitable sentence type for measuring language competence. As discussed, this may be because the jumbling type order is not appropriate; the arrangement influenced participants to apply the strategy of the first trial to the subsequent sentences, which affected the overall results. This suggests that a more controlled instruction in the design of jumbling order may be necessary, and this will be the focus of the next experiment.

The experiment has found that doing analysis from the most general perspective to a very specific and detailed level, and aggregating the pauses accordingly, has shown some improvement to the correlation values. The overall findings have found that (1) the range of correlations for each group has expanded, (2) the overall number of significant correlations has increased and (3) some data points have improved to a higher correlation. Some strong correlations and an effect of stimulus type order implies that (1) certain stimuli (e.g. *no jumbling*) can be used to measure grammar and vocabulary competence, and that (2) strategies of copying is a key issue to address. On top of that, this analysis also raises the questions of whether it is necessary to include the second and third sentence sets because it was found that the first sentence provides better data and that when participants did the second and third, they know what kind of sentence stimuli to expect. This is an interesting implication of the experiment which causes this study to rationalise the experiment design in the next experiment. This will focus only on two sentence sets, to confirm these findings.

On the other hand, the application of normalising has not been able to provide any clear pattern of improved correlations. This normalisation technique will be reconsidered in the next experiment, to confirm its usefulness.



## CHAPTER 7    Experiment 3: Investigating the Effects of Task Type (Processing Conditions) on the Strategy of Copying

### 7.1 Introduction

This final experiment aims to (1) investigate the effects of task types (processing conditions) given as instructions to control the processes of copying, in order to gain a better measure and find more consistent strategies, (2) identify the most effective jumbling condition to be used as a suitable sentence stimulus for measuring the cognitive processes in language processing, (3) confirm the effective measure of cognitive processes and finally, (4) test whether the normalisation technique is useful. This experiment builds upon the previous experiment, which showed that sentence order presentation affects the activity of copying, thus emphasizing the need to control the copying tasks, i.e. *task types*. The task types, otherwise known as processing conditions, are processing for meaning and fast-copying. With these two task types, participants were either asked to copy as quickly and as accurately as possible and at the same time attempt to understand the meaning of the sentence stimulus (i.e. processing by meaning), or to focus only on completing the copying as accurately as possible, without the requirement to understand the sentence (i.e. fast-copying). It is predicted that these task conditions will be able to exclude the influence of first trial copying over the subsequent sentence stimuli for all participants, and hence enable the distinguishing of the effective sentence stimuli.

Therefore, the first aim of this experiment is to observe the effects of each task type on the process of copying, in an attempt to control the effects of the sentence order. The two task types are predicted to control the copying process and eliminate the sentence order presentation effect, and as such should be able to produce strong correlations with ODT and VST. The newly-introduced task instruction should reduce the effects of the sentence order on the copying strategy. In order to confirm that participants understand what they copied, a set of True or False questions were asked immediately after the completion of copying each sentence. Questions were not asked for sentences that focused on fast-copying; with these manipulations, the effects of the aforementioned processing conditions were observed.

As a general prediction, processing for meaning should be able to distinguish different levels of nature of language processing more effectively than fast-copying. Processing for meaning requires

the participant to comprehend the sentence while copying it and their understanding is then assessed by their answering a True or False question. Fast-copying, on the other hand, forces participants to quickly copy without the need to understand, so the True or False question is not needed.

In the case of processing by meaning, depending on the type of jumbling received, it is predicted that participants who are of low competence and who are unfamiliar with the words in the sentence will (1) use smaller chunks, at word and syllable level, (2) have higher PLP values, and thus produce many long L2 and L1 pauses, (3) be unable to comprehend the sentence fully and therefore will be unable to correctly answer the True or False questions. Participants who are of high competence and are familiar with more of the words in the sentence will (1) have a larger chunk size, at word group and word level, (2) exhibit lower PLP values, and thus produce fewer long L2 and L1 pauses, (3) be able to comprehend the sentence better than the less competent participants and therefore correctly answer most of the True or False questions. Given these predictions, this experiment should help to differentiate levels of competence, through the task conditions.

The second objective is to identify the effective jumbling condition to be used as a sentence stimulus for measuring the cognitive processes of language processing. This objective follows the previous experimental work. To recap the previous exploration: Experiment 1 demonstrated that the effects of jumbling effectively differentiate varying levels of language processing. Experiment 2 extended this research, particularly with regard to the jumbling conditions. These results indicated, however, that a lack of control in sentence stimuli arrangement produces a sentence order presentation effect. The present experiment therefore repeats the same stimuli and jumbling conditions as used in Experiment 2, in order to find a way to deal with the sentence order presentation effect. With the new parameter of processing conditions (i.e. task types) introduced in this chapter, the participants' copying tasks can be controlled by instructing them to focus on either processing by meaning or fast-copying. This new instruction is hoped to assist the study in obtaining better data than in the previous experiment. Thus, with the ability to control the sentence order presentation effect on copying strategy, this chapter will be able to focus on identifying the most effective jumbling condition that can be used to measure the cognitive processes of language processing.

The third objective is to confirm the effective measure for cognitive processes in language processing. Following investigation in the pilot study, this research has concluded that the *median of pauses* (L1s and L2s) is the most effective statistical measure. This was applied in Experiments 1 and 2 and has been proven useful by the strong correlations produced. It is imperative, however, to revise the measure with more sophisticated sentence types, such as those to which the jumbling condition has been applied, in order to confirm its effectiveness. This chapter will take steps towards confirming the suitability of *median of pauses* in measuring the jumbling condition.

Finally, the normalisation technique will be applied, in order to observe whether this could compensate the pause data. This chapter repeats the same exploration on normalisation as used in previous experiments, to test its effectiveness in improving the data.

## 7.2 The Sentence Stimuli Design

The design of the stimuli is similar to that in Experiment 2, but is simplified to cater for the objectives outlined in Section 7.1.1.

### 7.2.1 The Baselines

The same baselines as were used in the previous experiments were adopted again, although in this experiment they focused on name writing (NW). Instructions to perform the NW baseline task were given at three different time intervals. The task was conducted under two different conditions: upper case and lower case. In each NW task, the name had to be written three times in order to obtain a mean result. The first NW task was conducted at the beginning of the experiment, before the copying activities started. The next took place after the completion of the first four sentences and the third was done at the end of the copying activities. Similar to in the previous experiments, the repetition of the NW activity is necessary in order to ensure the fair distribution of pauses across the baseline tasks, and thus calculate an average of the L2 (word level) and L1 (letter level) pauses.

### 7.2.2 The Sentences

The sentences used for the copying activities were selected from the 12 sentences used in Experiment 2 (Chapter 6). Due to the specific focus of this experiment, however, which is to examine the task types, the number of sentences was reduced to eight. All eight sentences, as

detailed in the previous chapter, had approximately the same level of word difficulty. It is predicted that if the processing conditions are effective, then the order of the stimuli or jumbling types will not have the effect on the subsequent sentences that was found in Experiment 2, where the order of the processing conditions or task types was presented to the participants in an alternating form.

### 7.2.3 The Processing (Task Type) Conditions

The two task type conditions introduced in this chapter, which aim to control the copying processes, are explained below:

a. Process of copying under the meaning condition.

This task is called *processing by meaning*. Participants are asked to copy the sentence stimuli, while simultaneously trying to understand the overall meaning of the sentence. This is similar to the 'immediate copying' approach, with more precise instructions given. Upon completion of the copying, participants have to answer a True or False question. Participants of all levels of competence will need extra time to complete the copying, as they simultaneously process the sentence for meaning. However, the levels of processing and pause lengths might differ depending on the familiarity level of the words in the stimuli. When the task has a jumbled condition, the level of difficulty in processing is increased. The difficulty in comprehending the sentence will be reflected in the answers provided for the True or False question. For example (refer to Figure 7.3), the sentence in the *both jumbled* condition (1d) is difficult to comprehend, so it is not easy to answer the True or False question. However, with a *no jumbling* condition, the sentence is more direct and requires less processing, making it easier to comprehend and to answer the True or False question. To conclude, the effect of this task type would require participants to try and comprehend the sentence in order to be able to answer the True or False question.

b. Process of copying under the fast-copying condition.

This task-type is called 'fast-copying'. Under this condition, participants are asked only to copy as quickly and as accurately as possible. They are not required to understand the meaning of the sentences. There are no True or False questions at the end of the task. This task type differs from the previous one because it does not test comprehension, but instead tests the level of familiarity

when copying quickly. Given that competent participants are familiar with the language, they would be able to produce short pauses during the copying, whilst less competent participants might take longer to copy. With this in mind, it is expected that the results will be different to those in the processing meaning task. To conclude, the effect of this task type should be able to distinguish the competence level based on the familiarity level of participants.

It is anticipated that introducing the two tasks and clarifying the instructions given for each task will be able to produce data from a consistent copying strategy.

## 7.3 Method

The present experiment was designed to determine whether the strategy of copying can be controlled by introducing specific instructions (i.e. the task types of meaning and fast-copying). In support of this, different measures were explored and different variables were combined and investigated in order to study the effects of each (see Figure 7.2).

### 7.3.1 Participants

Twenty-four Malaysian participants were recruited, aged between 20 and 35 years old. The same criteria for recruiting participants are used as used in Experiment 1; refer to Chapter 5 for more details.

### 7.3.2 Apparatus

The same apparatus as presented in Section 5.2.2 was used in this experiment.

### 7.3.3 Procedure

The participants were asked to copy the sentence stimuli given to them according to the task type assigned to them. The stimuli were visible at all times, but the participants were not allowed to read them in advance. The participants were only allowed to look at the stimuli and begin to write when the researcher said 'start'. Depending on the processing condition, or task type, the participants were required to process copying either by meaning or fast-copying (see Section 7.2.3). Processing by meaning was followed by a set of True or False questions, which needed to be answered after each stimulus. Fast-copying did not include any True or False questions and instructions. The name writing (NW) baseline task was conducted on three occasions: before the

copying task began, after the copying of the first four stimuli and following the completion of all eight stimuli. This approach is similar to the one applied in Experiment 2. During the copying activity, participants were trained to begin each sentence with a 'hash' (#) in order to ensure that the pause preceding the first stimulus letter could be validly measured. The participants copied 14 stimuli in the orders given in Figure 7.2. As participants were not presented with the same order of stimuli and jumbling types, the sentence order was unique for every participant.

		Participants Number																								
Stimuli		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	
	1	1a	3b	5c	5d	3a	2b	6c	1d	5a	2b	3c	8d	4a	2b	8c	4d	5a	3b	2c	1d	4a	8b	7c	7d	
	2	2b	2a	4b	8b	7c	3d	4d	3c	3d	8c	4a	2a	6c	3d	3a	5b	6b	1c	5b	6a	6d	3a	8d	4c	
	3	3c	6d	7a	4c	1b	4a	2a	5b	1c	5d	5b	6b	1d	5c	6d	6a	7d	8a	4d	3c	2b	6c	1b	1a	
	4	4d	1c	3d	6a	6d	5c	7b	7a	7b	3a	6d	4c	3b	1a	4b	7c	8c	5d	3a	7b	8c	1d	2a	6b	
	5	5a	5b	6c	1d	5a	6b	8c	2d	6a	4b	7c	1d	7a	6b	1c	8d	1a	4b	6c	2d	5a	4b	3c	2d	
	6	6b	4a	1b	7b	2c	7d	5d	4c	4d	6c	8a	7a	2c	7d	5a	1b	2b	7c	1b	4a	3d	7a	4d	5c	
	7	7c	7d	8a	2c	4b	8a	1a	8b	2c	7d	1b	3b	5d	4c	2d	2a	3d	2a	8d	5c	7b	2c	5b	8a	
	8	8d	8c	2d	3a	8d	1c	3b	6a	8b	1a	2d	5c	8b	8a	7b	3c	4c	6d	7a	8b	1c	5d	6a	3b	
LEGEND:																										
Task Type:																										
	Process by Meaning	x/	x is sentence number										a	No Jumbling - NWNL												
	Fast-copying	y	y is jumbling type										b	Words Jumbled - JWNL												
													c	Letters Jumbled - NWJL												
													d	Both Jumbled - JWJL												

Figure 7.1 Sentence order as given to the participants

In Figure 7.1, the darker boxes represent tasks for processing for meaning, whilst the white boxes show tasks for processing for fast-copying. Overall, there were 12 participants who started the experiment with task type meaning first and another 12 that started with task type fast-copying first.

SS	Q	SS	Q	SS	Q	SS	Q	SS	Q	SS	Q	SS	Q	SS	Q
1a	6	2a	6	3a	6	4a	6	5a	6	6a	6	7a	6	8a	6
1b	6	2b	6	3b	6	4b	6	5b	6	6b	6	7b	6	8b	6
1c	6	2c	6	3c	6	4c	6	5c	6	6c	6	7c	6	8c	6
1d	6	2d	6	3d	6	4d	6	5d	6	6d	6	7d	6	8d	6

LEGEND:			
SS	Sentence Stimuli	a	No Jumbling - NWNL
Q	Quantity	b	Words Jumbled - JWNL
		c	Letters Jumbled - NWJL
		d	Both Jumbled - JWJL

Figure 7.2 Equal distributions of sentence stimuli and jumbling type

Figure 7.2 shows an equal distribution of sentence stimuli and jumbling type, for the purpose of having a balanced amount of stimuli. The combination of numbers and letters represents the sentence number and the jumbling type. Each condition will appear six times across all participants.

NO	SENTENCES	Questions
1a	the representation of dramatic narrative in pictorial rather than in visual experiential terms is evident in images	Is the sentence about a 'drama festival'? <input type="radio"/> True <input type="radio"/> False
1b	than the in representation visual of experiential dramatic terms narrative is in evident pictorial in rather images	Is the sentence about a 'drama festival'? <input type="radio"/> True <input type="radio"/> False
1c	the retresenpation of dratamic natrarive in pirtocial rahter tahn in viusal expetienrial temrs is evinedt in imegas	Is the sentence about a 'drama festival'? <input type="radio"/> True <input type="radio"/> False
1d	tahn the in retresenpation viusal of expetienrial dratamic temrs natrarive is in evinedt pirtocial in rahter imegas	Is the sentence about a 'drama festival'? <input type="radio"/> True <input type="radio"/> False

Figure 7.3 Example of True or False questions given to participants

Figure 7.3 provides an example of a sentence with four different jumbling types and a True or False question (the same for each jumbling type) assigned. As a recap, the four jumbling types are *no jumbling* (a), *words jumbled* (b), *letters jumbled* (c) and *both jumbled* (d). Each sentence will have four different jumbling conditions, but, the same question is assigned to the four. Only one of these sentence conditions, with the True or False question, will be assigned to participants.

#### 7.3.4 The Language Tests

The same two independent tests were applied in this study as in previous experiments: the ODT and the VST. Descriptions of the tests can be found in Section 6.2.3 (VST), and in Sections 5.2.4.3 and 6.2.4 (ODT).

#### 7.3.5 Overall Structure of this Experiment: The Variables

Figure 7.4 shows the overall structure of this experiment.

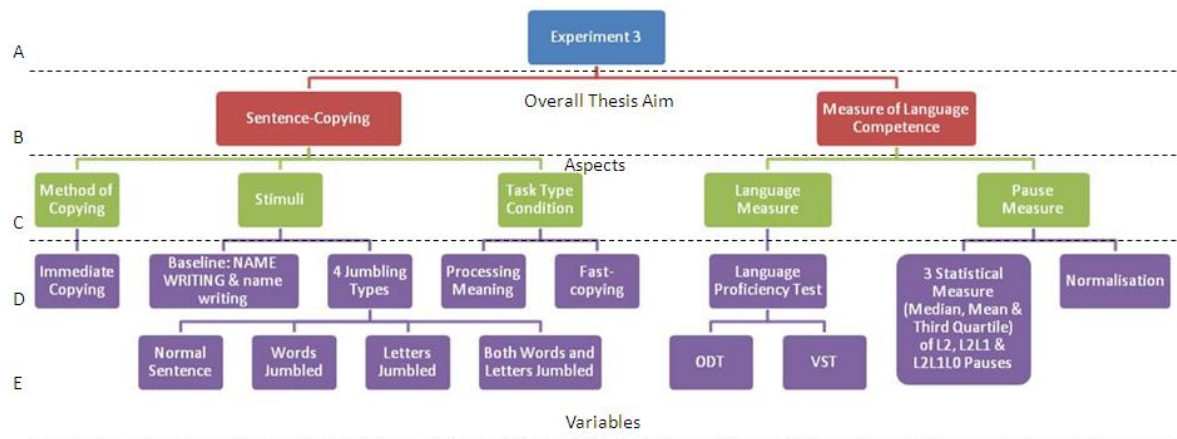


Figure 7.4 The overall structure of this experiment

Four aspects presented at level C group a number of components (levels D and E) that are important to this study. Some of these variables were established in previous experiments, and some were developed in this chapter. Variables used in previous experiments are (1) Immediate Copying (IC), (2) both tests of ODT and VST, and (3) normalisation. The additional variables introduced in this chapter are (1) the two task type conditions and (2) the exploration of three statistical measures, the median, mean and third quartile, with a combination of pauses, i.e. L2, L2L1 and L2L1L0. This experiment works from the same set of stimuli as Experiment 2, but only 8 of these stimuli are used, instead of all 12. The same name writing (NW) tasks are applied here as in previous experiments.

### 7.3.6 Analysis

The present experiment focused on investigating the effects of the newly-introduced processing condition (i.e. the task type) in controlling individual strategies that develop in order to facilitate the process of copying. To that end, this study has examined the contributions from (1) the independent tests of grammar and vocabulary, (2) the order of sentence presentation and jumbling conditions and (3) the order of the task types. Since the jumbling conditions have been shown to create the propensity to chunk at syllable level, only involving L2 pauses in the analysis was not considered to be sufficient. The findings suggest that the combination of L0, L1 and L2 pause levels should be used in the experiment. They also suggest a need for further investigation on the most suitable stimulus for measuring the cognitive processes of language processing. Using the combination of pause levels should capture the long pauses produced at syllable chunking as well as at word chunking. In support of this, the present experiment explores the Median, Mean



and Third Quartile (Q3), in order to calculate which most effectively differentiates the varying levels of cognitive processes.

## 7.4 Results

The results presented in this section demonstrate first the overall language scores, then the general effect of the task type (i.e. processing condition) on the different combination of pause levels and measures of cognitive processes in language processing.

### 7.4.1 General Language Competency Scores: ODT, VST & Pauses



Figure 7.5 Relationships between ODT and VST

Figure 7.5 shows the scatter plot graph of the data comparing the ODT and VST scores. The x-axis presents the ODT scores and the y-axis presents the VST scores. Both scores show a good variability of the participants' competency level: the ODT ranging from 33% to 75%, and the VST from 3800 to 11700. The correlation between the ODT and VST tests is weak,  $r(22) = 0.300$ . This was expected, as each test assesses a different language component, i.e. grammar (ODT) or vocabulary (VST).

The pauses are analysed at an aggregated level, using three different statistical measures and three different combinations of pause levels correlated against the Rank Score (Table 7.2). The L2 pauses show a consistent correlation across all measures: median, mean and third quartile (Q3).

When pauses are combined, e.g. L2L1 and L2L1L0, the mean and Q3 show a stronger correlation, however, the case is different with median.

	Median	Mean	Q3
L2	-0.386*	-0.370*	-0.364*
L2L1	-0.193	-0.370*	-0.430*
L2L1L0	-0.122	-0.407*	-0.420*

Table 7.1 Correlations performed with three measures and three combinations of pauses (\*p<.05)

#### 7.4.2 The True/False Questions

Four sets of questions were asked to each participant after completing the meaning condition task type, regardless of whether it was in the first or the second set. In total, 96 questions were asked, with 24 questions per jumbling type. The numbers of correct responses, according to jumbling type, are summarised in Table 7.7.

Jumbling Type	No of Correct Responses
a	19
b	19
c	18
d	13

Table 7.2 Number of correct responses based on the jumbling type received

Jumbling type 'd' was found to have the smallest number of correct responses. As shown in Table 7.8, the first set of True or False questions (in first position) has demonstrably fewer incorrect answers across all jumbling types than other positions. The fourth set (fourth position) has the highest number of incorrect answers, despite the amount of copying practice the participants had done. Segregating the correct responses based on the jumbling type did not make any difference as there was a fair distribution of wrong answers between types 'a' and 'b'.

	Position			
Jumbling Type	1st	2nd	3rd	4th
a	0	2	1	2
b	1	2	0	2
c	2	1	0	3
d	0	3	4	4

Table 7.3 The number of incorrect responses based on the jumbling type and the order in which they were received

To conclude, the first trial of any sentence type proved that participants were significantly more focused during that first trial. As the number of trials increased, participants became more prone to making mistakes, which might indicate that their focus has been affected by the jumbling type.

### 7.4.3 The Effects of Task Type (Meaning & Fast-copying) at General Level

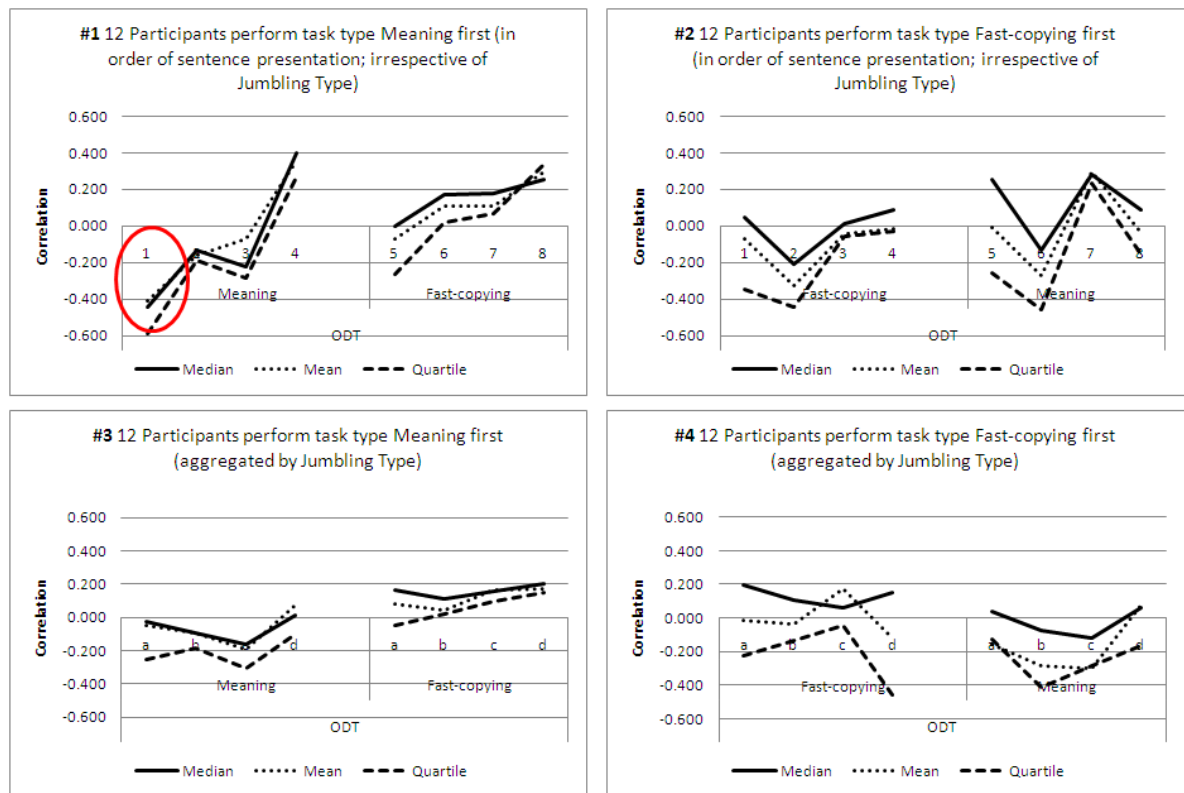
The purpose of the processing condition or task type (Meaning and Fast-copying) is to control the strategies of copying affected by the sentence order. The analyses performed have revealed that the jumbling condition applied in this experiment affects chunking at syllable level. Using the Median and only L2 pauses, as in the previous experiments, obscures processes taking place at syllable level, because L2 only considers pauses at the beginning of a word, right before the first letter. The decision to combine pause levels of L2L1 and L2L1L0 level was therefore taken, to discover the effects of syllable chunking on language processing. However, this chapter will only present data of the combination of 'L2 and L1' (i.e. L2L1) pauses, based on the assumption that participants would rarely chunk language meaningfully at stroke (L0) level. Moreover, detailed analysis carried out comparing the inclusion or exclusion of L0 has demonstrated that this does not make any noteworthy difference to the results. The findings from Experiment 2 show that effects of copying under the jumbling conditions occur more frequently during the first and second trial. The results here will therefore discuss the effects observed in the first trial.

As explained in Section 7.3.6, the results will be analysed in relation to (1) the independent tests, the ODT and the VST; (2) the order of sentence presentation (e.g. 1, 2, 3, and 4) or jumbling type (e.g. a, b, c and d); and (3) the order of task type (e.g. first the meaning followed by the fast-copying task type or vice versa). The jumbling types are named a, b, c, and d, representing *no jumbling*, *words jumbled*, *letters jumbled* and *both jumbled*, respectively. When data is aggregated at a general level, which includes all 24 participants, no strong correlation effects are observed. The analysis yielded inconclusive results on the presence of any particular effects during the implementation of each task type. When participants are segregated according to task type, however (meaning condition or fast-copying condition), the results begin to reveal interesting findings.

Figures 7.6 and 7.8 present the results for the ODT and VST tests respectively, grouped by task type. Of the four graphs set in each figure (#1, #2, #3 and #4), #1 and #2 present the data following the order of the sentence stimuli as given to the participants. #3 and #4 present the data

aggregated by jumbling type. #1 and #3 start with the meaning task type, while #2 and #4 begin with the fast-copying task type. Each line in the graphs represents the correlation results under three different measures: the Median, Mean and Q3.

Figure 7.6 shows the strongest correlation with the ODT during the first trial in graph #1:  $r(22) = -0.592$  ( $p < .01$ ), under conditions (1): Meaning task type appearing first; (2) following the sequence order of sentence stimuli presented to the participants; and (3) Q3 of pauses having the strongest correlation. More detailed correlation values for each graph in Figure 7.6 are presented in Appendix 7.1.



**Figure 7.6 Results showing the effects of Task Type Order and Sentence Presentation on ODT with L2L1 pause data**

To confirm the choice of pause level combinations, Figure 7.6 analyses and compares the correlation values of the first trial across the three different combinations of pauses. At L2 pause level, the median of pauses records the strongest correlation value,  $r(22) = -0.455$  ( $p < .05$ ). The L2 and median of pauses are the measures used in previous experiments. With L2L1 and L2L1L0, the Q3 of pauses shows the strongest correlation,  $r(22) = -0.592$  ( $p < .01$ ) and  $r(22) = -0.568$  ( $p < .01$ )

respectively. Similar findings observed between L2L1 and L2L1L0 demonstrate that focusing only on the combination of L2L1 is sufficient for this study, from here on in.

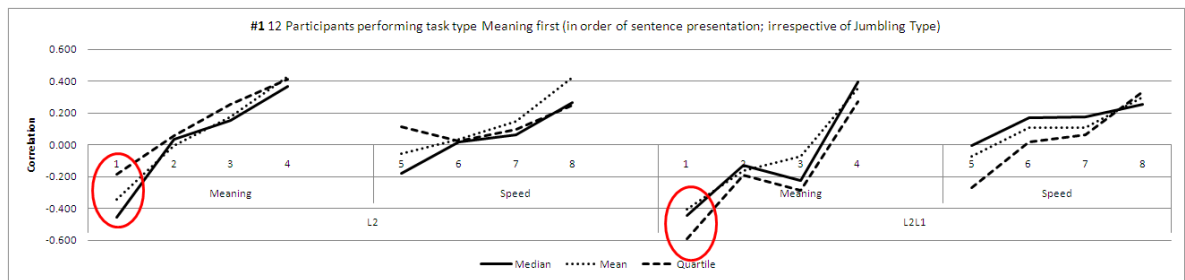


Figure 7.7 ODT Results at all pause levels according to the order of sentence presentation

The correlations with the VST scores present a different story. Figure 7.7 summarises the data, which demonstrate good correlation values in copying the sentence stimuli with most conditions.

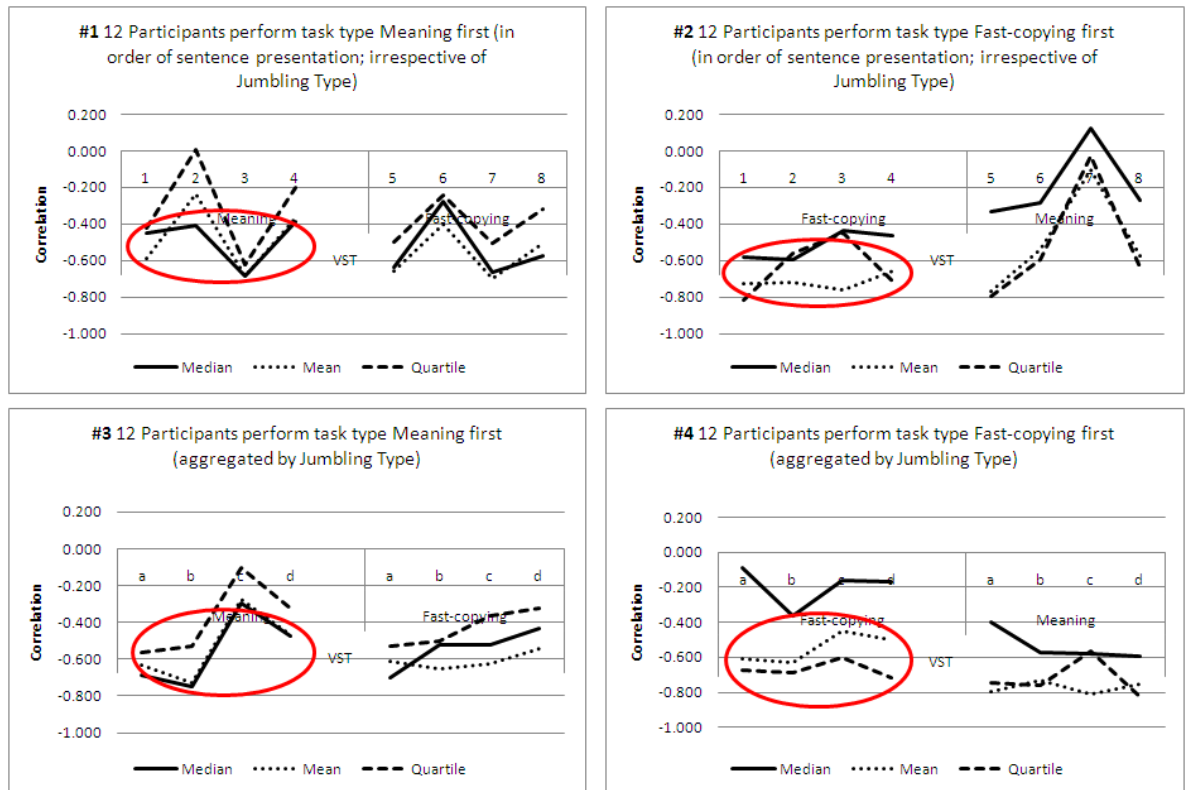


Figure 7.8 Results showing the Effects of Task Type Order and Sentence Presentation on VST with L2L1 pause data

#### 7.4.3.1 *Short Summary of the Results*

In search of the most suitable measure for the cognitive processes of language processing and the most effective sentence stimulus, strong correlations must be observed from both independent language tests (ODT and VST). The previous section, however, has only shown a strong correlation for the ODT during the first trial with Meaning as the task type and the order of sentences as presented to the participants, irrespective of the jumbling type. Below is a summary of the findings:

1. Implementing the Meaning condition first appears to be meaningful for the purpose of measuring language processes. The obligation that accompanies the instructions given to the participants to copy and process meaning at the same time helps to prevent the involvement of a personal strategy of copying. The participants, knowing that copying via meaning will be followed by a set of True or False questions, would focus entirely on the copying activity, as opposed to focusing on a copying strategy. Assuming that the effects of sentence type order are controlled, the task will be able to differentiate levels of language processing. The process of chunking and the effects on pause length production by competent and less competent language users is the same as that already explained in the previous experiments.

When participants are instructed to focus on fast-copying, without any obligation to understand meaning, they refrain from recalling or retrieving the meanings of words and focus instead on encoding and decoding the words as quickly as possible in order to complete copying successfully. Fast-copying assesses the immediate word knowledge of all participants, but also assumes that competent participants would be able to copy more fluently than the less competent.

This has been proven by the strong correlations apparent in Figures 7.6 and 7.8, for both ODT and VST, especially in the first trial and when the sentence stimuli are in the sequence order of presentation received by the participants.

2. Pauses at L2 level represent pauses at the beginning of the first letter of a word. The types of jumbling applied in this experiment manipulate and encourage chunking at syllable level (e.g. when processing long unfamiliar words or jumbled letters in words). Considering only the value of L2 pauses does not account for processes that occur within a word. This suggests the need for further exploration, focusing on the combination of

pause levels (L2L1 or L2L1L0). Assuming that processing difficult words or sentences will produce long pauses, a measure that could interpret these long pauses and locate differences between the varying competence levels is needed; hence the inclusion of the Mean and the Q3 of pauses.

L0, which reflects pauses at stroke level, can be excluded in this case on the grounds that participants rarely pause within a letter, and that it lacks meaning. Observations from the results have shown that there is only a slight difference in the respective correlation values.

3. ODT shows good correlations with the order of sentences presented to the participants, but not when grouped by jumbling type (at general level). VST, on the other hand, shows strong correlations with both conditions. In order to achieve the aim of this thesis, strong correlations must be demonstrated for both ODT and VST. Thus, further investigations were carried out on the first trial for both ODT and VST (Section 7.3.3). However, it is interesting to note that the ODT presents strong correlations only during the first trial, based on the order of sentences received, and this merits further investigation. The order of sentences presented contains a mixture of jumbling types, highlighting the need to carry out further analyses by segregating the jumbling types. This means that the data should be segregated further to a higher level, in order to discern the effects on both ODT and VST. Accordingly, results were divided into groups of six, depending on the jumbling type the participant received.

#### **7.4.4 The Effects of the Meaning Task Type during the First Trial**

In this section, the results focus only on the effects occurring during the first trial. It is predicted that the first trial will not be affected by the number of practice tasks, and would constitute the most precise dataset for further investigation. The analysis of the first trial is however restricted, due to the counterbalancing of jumbling type order (which is made unique to every individual). This restriction resulted in having processing by meaning to start with for sentences of jumbling type 'a' (*no jumbling*) and 'c' (*letters jumbled*), and fast-copying to start with for sentences of jumbling type 'b' (*words jumbled*) and 'd' (*both jumbled*). Each jumbling type consists of six datasets from different participants.

Figure 7.9 summarises the effects of correlations on the first trial only, segregated according to the jumbling type. Table 7.4 shows the correlation values in more detail.

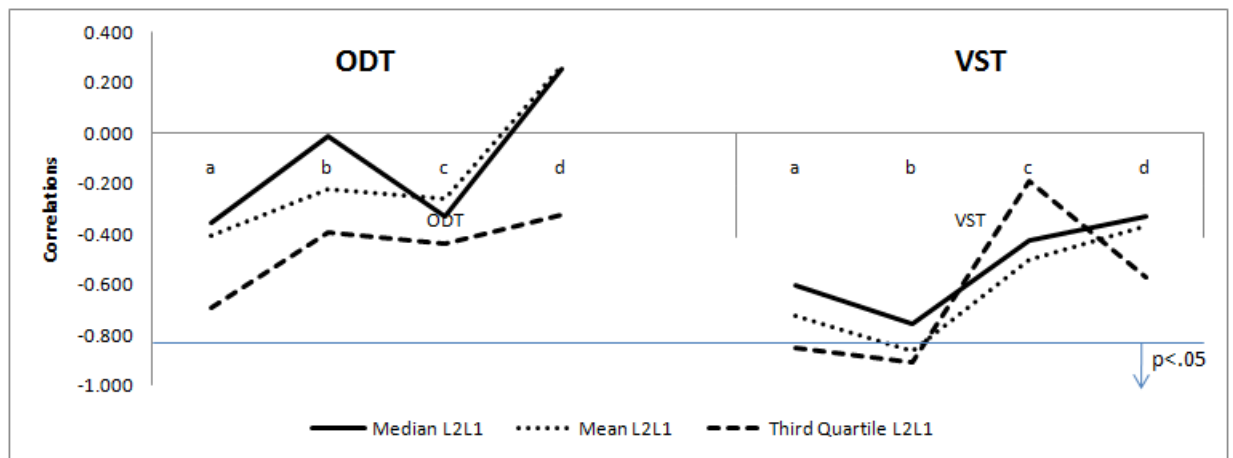


Figure 7.9 Correlations of the first trial (L2L1) for each jumbling type (6 participants each) against ODT-VST for each measure

L2L1	ODT				VST			
	a	b	c	d	a	b	c	d
Median	-0.357	-0.014	-0.327	0.256	-0.602	-0.753	-0.426	-0.330
Mean	-0.405	-0.222	-0.258	0.267	-0.722	-0.862*	-0.500	-0.370
Third Quartile	-0.691	-0.396	-0.435	-0.323	-0.854*	-0.908*	-0.188	-0.570

Table 7.4 Correlations of the first trial for ODT and VST with different measures, grouped according to jumbling type (\*<.05)

At L2 level, Median shows the strongest correlations, but, as previously discussed, using this might obscure processes that occur at other pause levels, such as L1, so might not be useful for this study. The Q3, on the other hand, shows strong correlations when pause levels are combined, especially for jumbling type 'a' (*no jumbling*). It is not surprising to see a weaker correlation for 'c' (*letters jumbled*), as the spelling process is disturbed. Jumbled letters can be confused with irregularly spelt words, which makes them harder to process. Instead of manipulating grammar and vocabulary-related processes, type 'c' (*letters jumbled*) diverts participants to focus on processes related to spellings. Types 'b' (*words jumbled*) and 'd' (*both jumbled*) provide different results with the fast-copying approach; they do not produce a good correlation with ODT, but do with VST.

To conclude, segregating the data to a finer level of detail yields stronger correlations, especially for ODT. The findings indicate that the Q3 with the combination of pause levels (L2L1 or L2L1L0)



might be a good measure of cognitive processes in language processing. The suitable sentence type for measuring grammar-related processes is sentence type 'a'. By contrast, word-related processes can involve types 'a' (*no jumbling*) and 'b' (*words jumbled*), under the proviso that they follow the Meaning condition (i.e. Task Type) in their first trial. The correlations observed for sentence types 'c' (*letters jumbled*) and 'd' (*both jumbled*), on the other hand, suggests that there might be effects caused by the disruption in the spelling process.

#### 7.4.5 Confirming the Effects of the Fast-copying Condition in the First Trial

Assuming that the participants followed the instructions exactly, performing the copying tasks as quickly and as accurately as possible, the effects of copying based on fast-copying would be purely based on copying rate, without any time spent on processing meaning. Competent participants would be able to copy a number of words at a time as a result of an automatic word recognition process. Less competent language users, on the other hand, would concentrate on completing the copying, potentially processing words individually with long L2 pauses. In order to test the authenticity of the pause data, the following analysis was conducted. The pauses were compared with the baseline name writing task data, as name-writing is an automatic process with no meaning-related processing, because personal names are the most common words that an individual ever writes.

It is predicted that:

- Correlations between the fast-copying condition in the first trial and NW will display significant values, especially for types 'a' and 'b'. Copying under the fast-copying condition does not require the understanding of meaning, thus rendering the task akin to the name writing process, producing strong correlations. These results can be seen in Table 7.5.

CORRELATIONS	Jumbling Type			
	a – no jumbling	b – words jumbled	c – letters jumbled	d – both jumbled
NAME WRITING	0.720	0.606	0.305	0.482
name writing	0.726	0.751	0.185	0.304

Table 7.5 Correlation values between the fast-copying condition at first trial and NW (N=6, N.S)

- Correlations between the meaning condition in the first trial and NW will not produce good correlation values, primarily because name writing is not a meaning-driven activity. These results can be found in Table 7.6.

CORRELATIONS	Jumbling Type			
	a – no jumbling	b – words jumbled	c – letters jumbled	d – both jumbled
NAME WRITING	0.088	0.050	0.026	0.078
name writing	-0.066	-0.054	-0.036	0.054

Table 7.6 Correlation values between the meaning condition at first trial and NW

Tables 7.5 and 7.6 both agree with the predictions already made. It can be seen that correlations between jumbling types ‘a’ (*no jumbling*) and ‘b’ (*words jumbled*) and NW in both upper and lower cases are strongly significant with fast-copying as the first trial, but not with the meaning condition.

#### 7.4.6 Findings on Normalisation

Normalisation is a process for reducing any unwanted effects that occur during the process of writing or copying. The aim of this section is to examine whether the application of the normalisation technique improves correlation results and whether it is needed. In order to do so, data based on the Q3 measure of L2 and L1 (to be called as Q3L2L1) pause combinations were used. The original pause values are compared against normalised data with (1) L2L1 pauses during NW in upper case letters and (2) L2L1 pauses during NW in lower case letters. Normalisation however, could not be performed with L1 sentence stimuli values because the Q3L2L1 already includes the L1s of the sentence. Figure 7.10 presents the original and normalised data for comparison. The darkest line, which represents the Q3L2L1, maintains its strong correlation value. This suggests that normalising with the NW baseline task does not improve the correlation, especially for jumbling type ‘a’ (*no jumbling*) and ‘b’ (*words jumbled*).

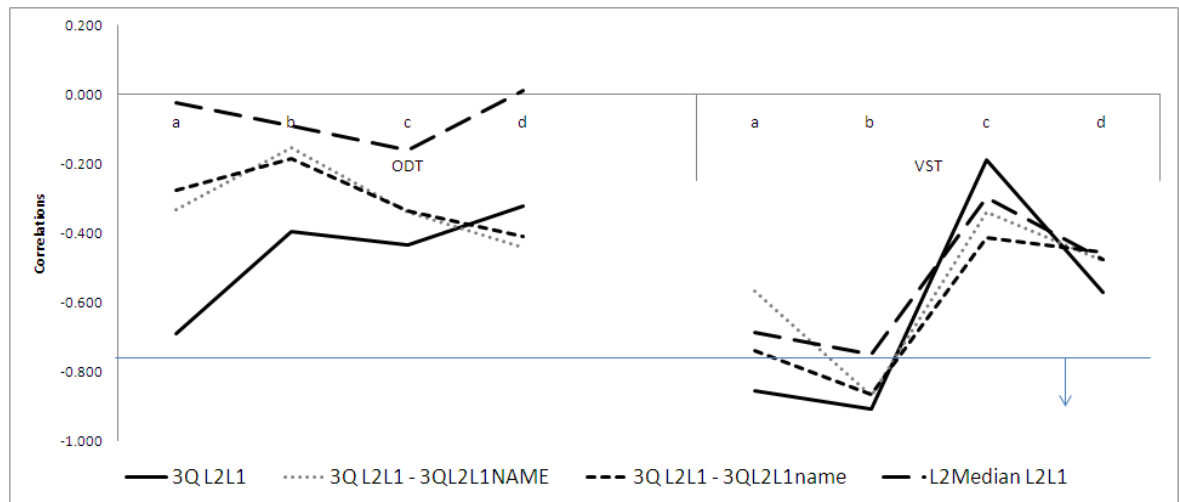


Figure 7.10 Normalisation conducted only on the First Trial

## 7.5 Discussion

This chapter introduces Task Type as an approach to control individual copying strategy that is affected by the *sentence order*, which was identified in the previous experiment. The Task Type manipulation requires participants either to process sentences through meaning first, followed by fast-copying, or vice versa.

### 7.5.1 Did the Task Type – Meaning or Fast-copying – Control Copying Strategy?

The order of the implementation of the Meaning or Fast-copying condition (first or second) affected the way participants processed copying. The Meaning condition was followed by a set of True or False questions, which were related to the sentence copied; the Fast-copying condition did not. The set of questions was to test whether participants understood the sentence they had copied. As predicted, it was found that introducing the Meaning condition first was useful in assessing the language comprehension of participants, and was more effective when introduced first, rather than after the fast-copying condition. The fast-copying condition, however, proved a good measure for the fluency in writing words.

In order to measure the components of cognitive processes in language processing, correlations must be strong when measured against both the independent measures, ODT and VST. In the case of ODT, strong correlations were only observed during the first trial, when the condition was segregated according to the order of sentences presented to the participants, irrespective of the jumbling type (see Section 7.4.2). VST, on the other hand, seemed not to be sensitive to any

measures or conditions and produced many strong correlations. This indicates that word-related processes occur at almost all stages, irrespective of the jumbling condition. When the analysis (for VST) was segregated further by jumbling type, still focusing only on the first trial, the correlation showed stronger values (see Section 7.4.3). This means that the participants' competence levels were revealed most clearly during the first trial, irrespective of the jumbling type received. Drawing on this finding, it can be argued that the introduction of the Task Type manipulation can control the copying strategies employed by participants.

Further analysis was conducted on the effects of the fast-copying condition. It was predicted that if participants faithfully followed the instructions for fast-copying, the processing costs would involve copying word-by-word (a word-related process but not a grammar-related process), and correlations with NW would therefore be strong. Name writing is considered an automatic process because of its highly practiced nature and the fact that it is well established in memory. As such, it would not require extra processes to digest meaning. When NW is correlated with the meaning condition pause data, however, weak correlations are expected because of the different processes involved. Copying in the meaning condition involves the recall and retrieval of information from memory, whereas NW can be automatically processed. These assumptions were proven correct, as explained in Section 7.4.4, especially for jumbling types 'a' and 'b', i.e. when words have correct spellings and are meaningful, because findings have shown that participants require a longer time to process meaning.

### 7.5.2 What is the Effective Measure?

Previous experiments have indicated that using the Median of L2 pauses was an effective measure. Using the Median is straightforward. It produces good correlations, which show significant results as required. In this chapter, however, where the jumbling criteria were applied for the sentence stimuli, the Median of L2 pauses did not expose all the processes occurring within a word (L1 letter level). Therefore, this chapter explored the Mean and Q3 measures, which were considered to be better suited for studying the combinations of different pause levels; L2L1 and L2L1L0.

A general analysis has found that with the ODT, a strong correlation can only be observed during the first trial under the following conditions: (1) when sentences are in the order of sequence presented to participants, (2) when the Q3 measure is used and (3) when either L2L1 and L2L1L0

pauses are measured. Findings were also observed to be stronger when the analysis focused on a finer level of detail: (1) looking only at the first trial and (2) grouping by jumbling type. Similar results were also observed with the VST. To summarise, the present experiment has concluded that when using stimuli manipulated under the jumbling conditions, the application of the Q3 statistical measure over a combination of pauses, especially L2L1 pauses, is able to distinguish the varying competence levels.

### 7.5.3 What is a Good Stimulus?

The four types of jumbling used in the experiments were intended to manipulate the difficulty of copying. The *no jumbling* condition is considered the easiest, and *both jumbled* the most difficult. *No jumbling* may contain unfamiliar technical terms, which can increase the level of difficulty for different levels of competence. The above findings show (1) *no jumbling* and (2) *words jumbled* to be the most suitable conditions for the sentence stimuli, for measuring grammar and vocabulary competence, as the correlations tend to improve with these two types. However, the findings are still not strong enough to confirm this, because of the limited number of participants, and this therefore requires further exploration. The *no jumbling* and *words jumbled* conditions preserve the meaning, even though *words jumbled* would be more difficult for less competent language users to comprehend, on account of the lack of grammatical structure. Both are, however, better measures than *letters jumbled*, which interferes with other processes, such as spelling knowledge. At some point, the copying of jumbled letters becomes meaningless to a less competent participant, especially when the words are unfamiliar and difficult to recognise. The *both jumbled* condition is useless for the purposes of measuring the cognitive processes in language processing, and can hence be dismissed.

The order in which the stimuli were given was unique to each participant. Participants did not receive the same sequence of jumbling types, based on the assumption that counterbalancing should not negate any sentence order presentation effect, if the task type instructions worked well. This differs from the previous experiment, where jumbling types were designed and allocated based on the Latin square procedure and did not include any specific instructions, such as in the task types introduced here. With the task type condition introduced and the counterbalancing of the jumbled stimuli sequence, the effects of individual copying strategies could now be controlled. For this reason, the effect of each jumbling type is dependent on the

task type condition given. When correlating these effects with the independent tests, the VST shows a higher influence on the jumbling types. It exhibits good correlations with both meaning and fast-copying conditions. The ODT, on the other hand, only displays good correlations with the meaning condition presented first, and only during the first trial, irrespective of jumbling type.

When participants were asked to copy under the Meaning condition first, the VST showed a strong correlation in the *no jumbling* and *words jumbled* conditions. Both the *no jumbling* and *words jumbled* used normal words. The *words jumbled* condition affects the grammatical structure by transposing the original word positions, and thus affects the processing of its overall meaning. The *letters jumbled* condition is more difficult than the *words jumbled condition* because the spelling manipulation inhibits the automatic recognition of words in the case of less competent participants. The *letters jumbled* condition forces participants to switch their focus from the meaning of the sentence to the correct spelling of the word. This does not test the nature of language processing in general, but instead would be suitable for researchers who are interested specifically in testing word recognition. The *both jumbled* condition is the most difficult. If it appeared on the first trial, it would take longer for participants to complete the copying. If it appeared at the end, i.e. in the fourth trial, pause lengths are a bit shorter because of the ‘practice effects’. When participants started with the fast-copying task type first, however, the copying of *no jumbling* and *words jumbled* conditions were much faster than with the meaning condition, as predicted.

#### 7.5.4 What is the Suitability of the True or False Questions?

The introduction of the True or False questions seems able to support the processing by meaning task. However, there is a need to amend the questions. The current sets of questions were found to be too direct, and require participants to process only certain words in order to be able to answer the questions successfully. This may have affected the copying of the rest of the sentences, as participants could have guessed the kind of strategy needed in order to be able to answer correctly. Therefore, more exploration is recommended in order to refine these questions, to make them more solid and able to assess whether participants have understood the meaning of the sentences properly, instead of allowing them to simply guess the correct answer.

### 7.5.5 What is the Effect of Normalisation?

Earlier experiments have found that normalisation assisted in the production of stronger correlations. Normalisation was also helpful in correcting the effects of fast-copying in writing, as described in Experiment 1. This technique was also used to refine the sentence stimuli in search of the most effective sentences for measuring the cognitive processes in language processing. In Experiment 2, normalisation assisted in improving the correlation values to study the effects of jumbling conditions on copying and language processing. In the present experiment, normalisation appeared to be embedded within the newly-introduced measure: the Q3L2L1. As presented in Section 7.4.6, this new measure is able to sustain the strong correlations, without the need to perform normalisation using NW. Even L2Median of L2L1 data showed weaker correlations compared to Q3L2L1. This finding not only points to a significant discovery, but also has made the measurement of language processing simpler.

### 7.5.6 Implications of the MoC

In this experiment, the instructions for copying are made stricter, to control the processes involved, by asking participants to focus on either processing by meaning or fast copying. When the instruction for copying is clearer, compared to the instructions in the previous experiments where the copying process was not controlled, it assists participants to focus better, in that they could control their copying in order to achieve what they are asked to do. For example, when copying requires them to process by meaning, their copying routes might involve semantic processing if they are unfamiliar with the words, or they might use a direct route if they are familiar with the group of words or individual words. However, when they are asked to do fast copying, they might completely avoid referring to semantic processing (because it is not required) and instead use the same route over and over again until copying is completed. The repeated route used depends on the level of familiarity. For example, if the copier is familiar with the words at group of letters level, i.e. syllable level, then they would repeatedly copy groups of letters, using the same route, until the word is complete. The same scenario occurs for word or group of words level.

The two new instructions were able to show two different perspectives of processes' effects on the pauses. Although we could easily understand the meaning of the instructions by the description of its name (Processing Meaning and Fast Copying) we do not however know the

detailed effects of each instruction on the copying; these can only be seen on the pause pattern. Detail analysis in this experiment, however, was able to show that these two instructions were able to control the underlying processes of copying, which altered depending on whether processing by meaning was needed.

## 7.6 Conclusion

The overall approach to data analysis showed stronger correlations when this was conducted with the top down analysis approach, from the most aggregate level down to the most segregated. In other words, correlations were stronger when the analysis focused on individual trials at a more detailed level. Rather than using Median of L2 (word level) pauses, this study has proposed the use of Q3L2L1 as a more suitable option. A summary of the contributions made in this chapter is below:

The choice of the effective measure for the cognitive processes of language processing is significantly related to the kind of stimuli applied in the study. For example, if the stimulus used difficult words, syllable chunking might occur, which would force the measure to include L1 pauses in the analysis. If the stimulus assessed word knowledge or recognition, it might need to include only L2 pauses, in order to measure language processing at word level.

The *no jumbling* (a) and *words jumbled* (b) conditions were found to be useful in assessing grammar knowledge because the condition of the sentences helps participants to process meaning. The *words jumbled* (b) condition was proven to be specifically useful for assessing word knowledge, because of the condition that made the words unconnected. The *letters jumbled* (c) condition seems to be good at assessing spelling knowledge, because it requires participants' recognition of the spellings in order to copy the sentences. The letters jumbled condition increases the difficulty of recognising words, but by being familiar with the correct spellings, participants can recognise the words and so copy more quickly.

Conducting the meaning task type first, followed by the fast-copying task type, was useful for assessing language processing. The fast-copying condition operated as a means to assess the fluency in copying words. When the fast-copying condition was conducted first, however, the



effects of processing meaning were hindered, because of the task effects resulting from the first four trials.

In terms of the experiment's limitations, the alternation of task type conditions, as given to the participants, resulted in an imbalance in the allocation of jumbling types. Instead of having an equal number of distributions, the experiment had 12 participants doing the meaning task type first, with the *no jumbling* (a) and *letters jumbled* (c) conditions, and another 12 participants performing the fast-copying task type first with the *words jumbled* (b) and *both jumbled* (d) conditions. A further limitation is the number of participants, six in each jumbling group, which prevents further analysis. Despite this imbalance, the results were able to show significant correlations and managed to achieve the objective substantial to this study.

The experiment, also implicated a number of design options which are considered to be beneficial for future research. It may be useful to:

1. Refine the set of True or False questions that come after the meaning task type, in order to confirm that the questions are asking about the understanding of the sentence.
2. Remind participants frequently of the instructions before each copying task.
3. Exclude jumbling types *letters jumbled* (c) and *both jumbled* (d), if the research objective is the assessment of language comprehension, because these two jumbling conditions may not serve the purpose of measuring the two language components. Type 'c', however, might be useful for assessing word recognition.
4. Exclude the application of the normalisation technique, as this might not be necessary, as the newly found measure, Q3L2L1 already includes the normalisation technique, embedded within the measure.
5. Examine copying on printed lines. Findings at the later stage, where the Q3 with pauses at all levels (L2L1L0) is applied, raise the question of whether copying letter by letter in a box would be necessary. It might be a good suggestion for future research to run a comparison between writing in boxes and printed lines, on the condition that letters are individually written; i.e. not cursive or joined-up handwriting.

As a conclusion, the experiment conducted in this chapter demonstrated that the introduction of the task type conditions was able to control the strategies of copying under certain circumstances. The findings indicate that, within the context of the stimuli used for this experiment, the Q3L2L1 measure was the most suitable measure of cognitive processes in language processing. Analysing the first trial of each jumbling type indicates the ability of pause analysis to measure the language components of grammar and vocabulary, as can be seen with the correlations with ODT and VST. In addition, the results were able to confirm the effectiveness of jumbling types 'a' and 'b' for measuring language processing through the activity of copying.

The next chapter will revisit the previous experiments in order to apply the measure identified in this chapter.

## CHAPTER 8 The Experiments Revisited

### 8.1 Introduction

In Experiment 3, a new measure, the Third Quartile of L2 and L1 (Q3L2L1), was found to be more effective than the Median of L2s (L2Median), with reference to the implementation of the jumbling conditions. However, it is still not clear if, out of these two measures for language competence; Q3L2L1 is the most suitable measure. Hence, in the attempt to validate and confirm this new measure this chapter revisits Experiments 1 and 2, applying the new measure.

To recap, Experiment 1 and 2 differ in terms of the sentences used: Experiment 1 explores several different kinds of sentences, while Experiment 2 focuses on academic sentences from journals. The pilot study itself is not involved in this investigation because of its limited number of participants and their varied language backgrounds. In this chapter, L2Median will be known as Method A, whilst Q3L2L1 will be known as Method B.

It is hypothesised that Method B will be the better measure of language competence, because it is capable of picking up long L1 (letter) pauses, which may be associated with syllables. Including syllable level pauses in the measure provides more information about the processes occurring within a word than using only L2 (word) pauses, which only takes into account the processing occurring at the beginning of a word.

### 8.2 Experiment 1 Revisited

First, Experiment 1 is revisited, and the results at the highest level of aggregation are compared, using Methods A and B. The investigation was furthered by looking at more detailed data – at the correlation values for each individual sentence – to analyse the correlation differences between the two measures. As a recap, Experiment 1 focuses on finding a suitable sentence stimulus to measure the cognitive processes of language processing.

## 8.2.1 Results

### 8.2.1.1 Comparing the Two Measures in General

In order to test whether it is useful to examine the value of Method B, a general analysis was performed by comparing Method B correlations values with those of Method A. The correlations were done against the ODT tests scores, and the results are shown in Table 8.1.

ODT Level	Method A	Method B
Basic	-0.585**	-0.496*
Intermediate	-0.681***	-0.616**
Advanced	-0.649**	-0.496*
Mean of Score	-0.687***	-0.577**

Table 8.1 Comparison of correlations between pauses and language performance using Methods A and B (\* $p < .05$ ; \*\* $p < .01$ ; and \*\*\* $p < .001$ )

Overall, the results show significant correlations for both methods. The correlations of Method B are not particularly as strong as for Method A, but are still significant. This suggests the need for further investigation comparing the two methods at a more detailed level: the results for each individual sentence.

### 8.2.1.2 Comparing the Two Measures in Sentence Analysis

Table 8.2 compares the correlation results of the two measures, and shows the means of pause values for each sentence for L1 pauses, L2 pauses and combinations of L2 and L1 pauses across all participants. These means of pauses show the average of median of pauses at different levels (i.e. L1, L2 and L2L1) across all participants for each individual sentence. The mean of pauses can indicate the difficulty of copying the individual sentence at letter or word level, by comparing the mean for an individual sentence to the overall mean of all sentences.

Set of Sentence Processing	Sentence Group	Sentence No	Sentence Description	Correlation Values		Pauses					
				Method A: (L2Median)	Method B: (L2L1 Third Quartile)	L1 only		Method A		Method B	
						Mean of Median of L1 Pauses (ms)	Mean of Median of L1 Pauses (ms)	Mean of Median of L2 Pauses (ms)	Mean of Median of L2 Pauses (ms)	Mean of Third Quartile of L2L1 Pauses (ms)	Mean of Third Quartile of L2L1 Pauses (ms)
<b>III</b>	<b>I</b>	24	Low frequency words, with no spaces in between. The last and first letters of successive words are high frequency bigrams.	-0.178	-0.660**	206.84		302.13		523.51	
	<b>I</b>	25	Low frequency words, with no spaces in between. The last and first letters of successive words are high frequency bigrams.	-0.100	-0.623**	347.70	277.27	508.60	405.37	543.01	533.26
<b>II</b>	<b>G</b>	21	Two proverbs combined by alternately selecting words	-0.691***	-0.607**	197.30		255.93		389.34	
	<b>D</b>	13	Words with jumbled letters, except for the end-letters	-0.660**	-0.529**	241.45		514.45		482.70	
	<b>H</b>	22	Three proverbs combined by alternately selecting words	-0.618**	-0.520**	208.95		267.85		364.58	
	<b>D</b>	14	Words with jumbled letters, except for the end-letters	-0.666**	-0.490*	241.78		532.05		471.78	
	<b>B</b>	9	Technical sentences with specialist terms from specific fields and no spaces in between words	-0.345	-0.478*	234.00		473.95		508.41	
	<b>B</b>	10	Technical sentences with specialist terms from specific fields and no spaces in between words	-0.664**	-0.477*	239.43		424.53		464.15	
	<b>E</b>	17	Garden path sentence	-0.702***	-0.467*	246.75		484.00		378.48	
	<b>A</b>	7	Technical sentences with specialist terms from specific fields	-0.416*	-0.460*	245.73		442.15		419.51	
	<b>H</b>	23	Three proverbs combined by alternately selecting words	-0.556**	-0.410*	192.18	227.51	244.95	404.43	380.45	428.82
<b>I</b>	<b>F</b>	19	Common words, with no spaces in between. The last and first letters of successive words are high frequency bigrams	-0.618**	-0.371	204.65		255.15		435.80	
	<b>E</b>	15	Garden path sentence	-0.684***	-0.339	253.83		539.48		369.39	
	<b>A</b>	8	Technical sentences with specialist terms from specific fields	-0.619**	-0.332	237.93		596.83		383.91	
	<b>E</b>	16	Garden path sentence	-0.584**	-0.318	241.40		531.58		406.03	
	<b>F</b>	18	Common words, with no spaces in between. The last and first letters of successive words are high frequency bigrams	-0.572**	-0.318	240.28		504.30		415.29	
	<b>G</b>	20	Two proverbs combined by alternately selecting words	-0.567**	-0.297	216.53		285.05		389.44	
	<b>C</b>	11	A proverb with jumbled words	-0.550**	-0.235	262.48		514.13		373.81	
	<b>C</b>	12	A proverb with jumbled words	-0.398*	-0.203	240.55	237.20	473.50	462.50	356.88	391.32
MEAN:						236.83	247.33	428.98	424.10	424.02	451.13

**Table 8.2 Correlations of Method A, Method B and Means of Pauses at different levels (\*p<.05; \*\*p<.01; \*\*\*p<.001)**

The results presented in the table are ranked according to the Method B correlation values, from the strongest negative correlation to the weakest negative correlation. The Sentence Group column provides information on the sentence type, as explained in Chapter 4. The Sentence Number gives a reference to the order of sentence presentation given to the participants. The number starts from 7 (upon the completion of 6 baseline writing tasks) and runs to 25, making a total of 19 sentences. The description of the sentences used is given in the Sentence Description column. The table also presents the mean of pause values for Median of L1s, Median of L2s and Third Quartile of L2 and L1. It was expected that the mean pauses for Method B would be higher than Method A, and be able to show the different nature of processing.

It was found that the changes in correlation values and the means of medians from Method A to Method B could be categorised into three groups. These are defined as:

Set I (bottom): correlations have decreased from being significant with Method A to being non-significant with Method B. The pause means of the sentences shows that the pause mean for Method B is smaller than Method A for 6 out of 8 sentences in the group. This highlights that this group has a very little occurrence of L1 syllable chunking.

Set II (middle): correlations remained approximately the same. The pause means of the sentences shows some evidence of syllable chunking. 5 out of 9 sentences have larger pause means in Method B than Method A, which is nearly equally balanced.

Set III (top): correlations have improved from being among the lowest (not significant) for Method A to being the very highest (strongly significant) for Method B. There is an increase in pause means for both sentences in the group.

The most interesting findings are the changes in correlation values seen with Set III after the implementation of Method B, providing results which may be useful for this study. It is suspected that there might be some effect from the L1 pauses, showing participants are copying groups of letters together (i.e. syllable processing). The impact is not only seen with Set III, but also provides the ability to group similar ways of processing together. From observations, the sentences that are grouped together have similar characteristics. For example, sentences in Set I contain grammars and words that can be easily recognised and understood, sentences in Set II are a slightly more difficult than Set I, in that there may be a different spelling version of a word (i.e. *letters jumbled*) or the grammar structure may have been disturbed (i.e. *words jumbled*), and sentences in Set III are the most difficult, because it is not easy to recognise the words when they are mixed up using high frequency bigrams at end-letters, and no spaces between words.

A pause distribution example for Set I:

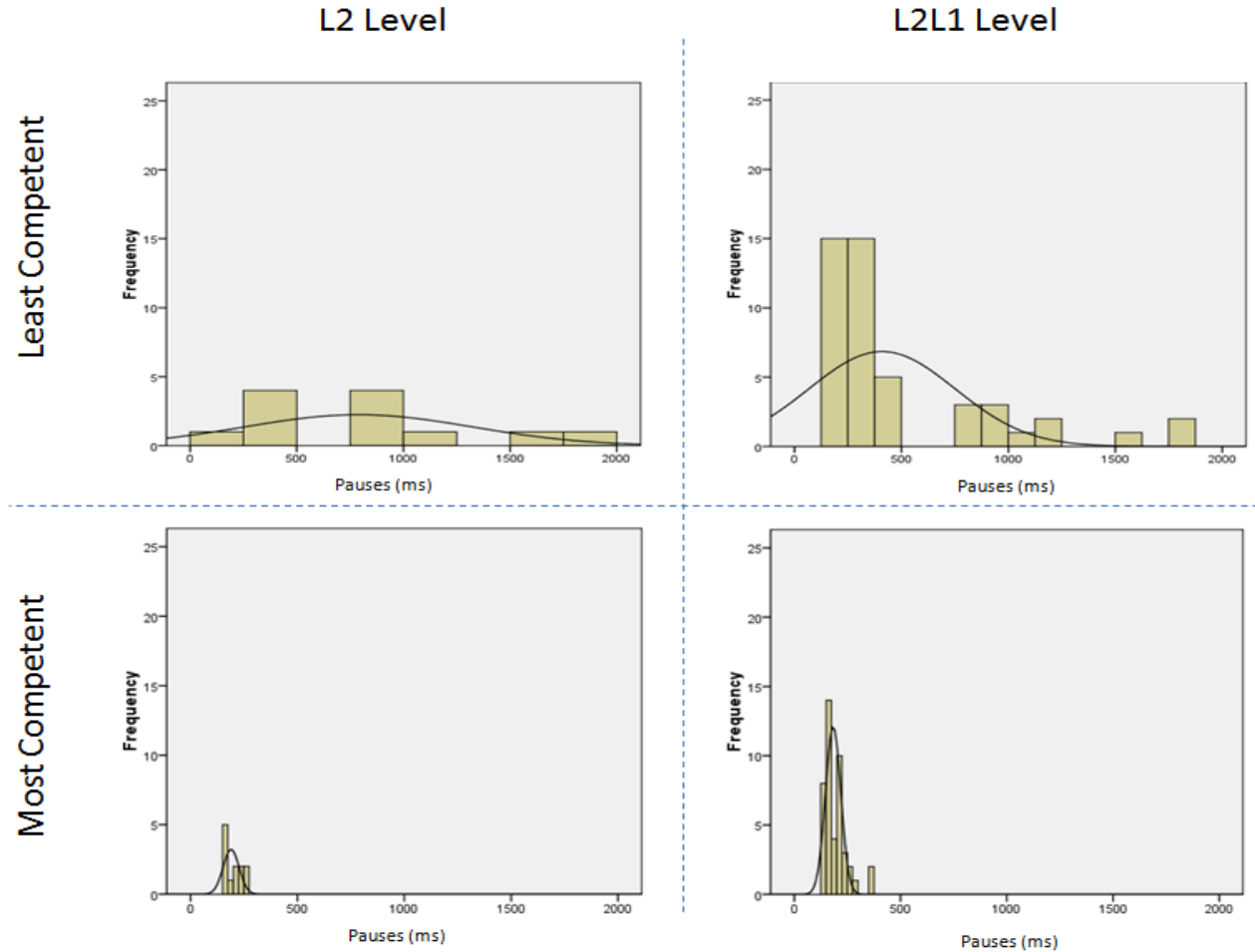


Figure 8.1 An example of pause distribution for Set I

Figure 8.1 shows the pause distribution for sentence 19, **'itwouldbegoodformotherthoughithinkitcanbebetter'**. It shows the pause distribution for the most and the least competent participants, for each different method. The figure shows that for the most competent (MC) participant, both L2 Level and L2L1 Level pauses are clustered at the very low values. This shows that MC was able to process the copying of this sentence fluently, without the need for long pauses to process unfamiliar words within this sentence. For the least competent (LC) participant, the pause values are much more widely distributed, but with a stronger skew towards the higher values for L2L1 Level. This indicates that LC had difficulties recognising the words because of the omission of spaces, hence producing longer pauses at both L2 and L2L1 Level.

A pause distribution example for Set II:

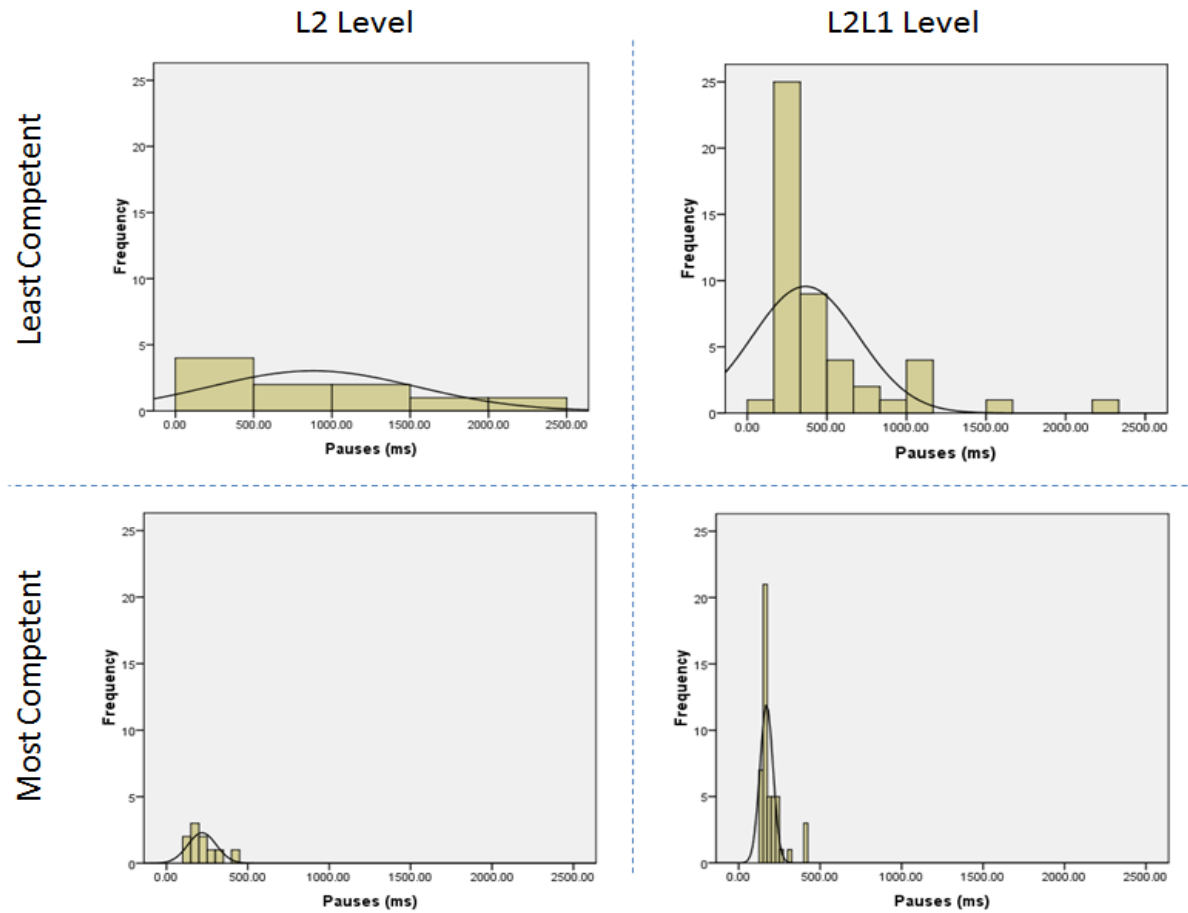


Figure 8.2 An example of pause distribution for Set II

Figure 8.2 shows the pause distribution for sentence 21, '**all actions that speak glitters louder is than not words gold**'. Here again, MC shows a strong cluster for both L2 and L2L1 Levels at low pauses, and LC's pause values are much more widely distributed, with a strong skew at the higher values of L2L1 Level. The difference between this Set II and Set I is that Set II has a much higher frequency of pauses at a lower value (below 500ms).



A pause distribution example for Set III:

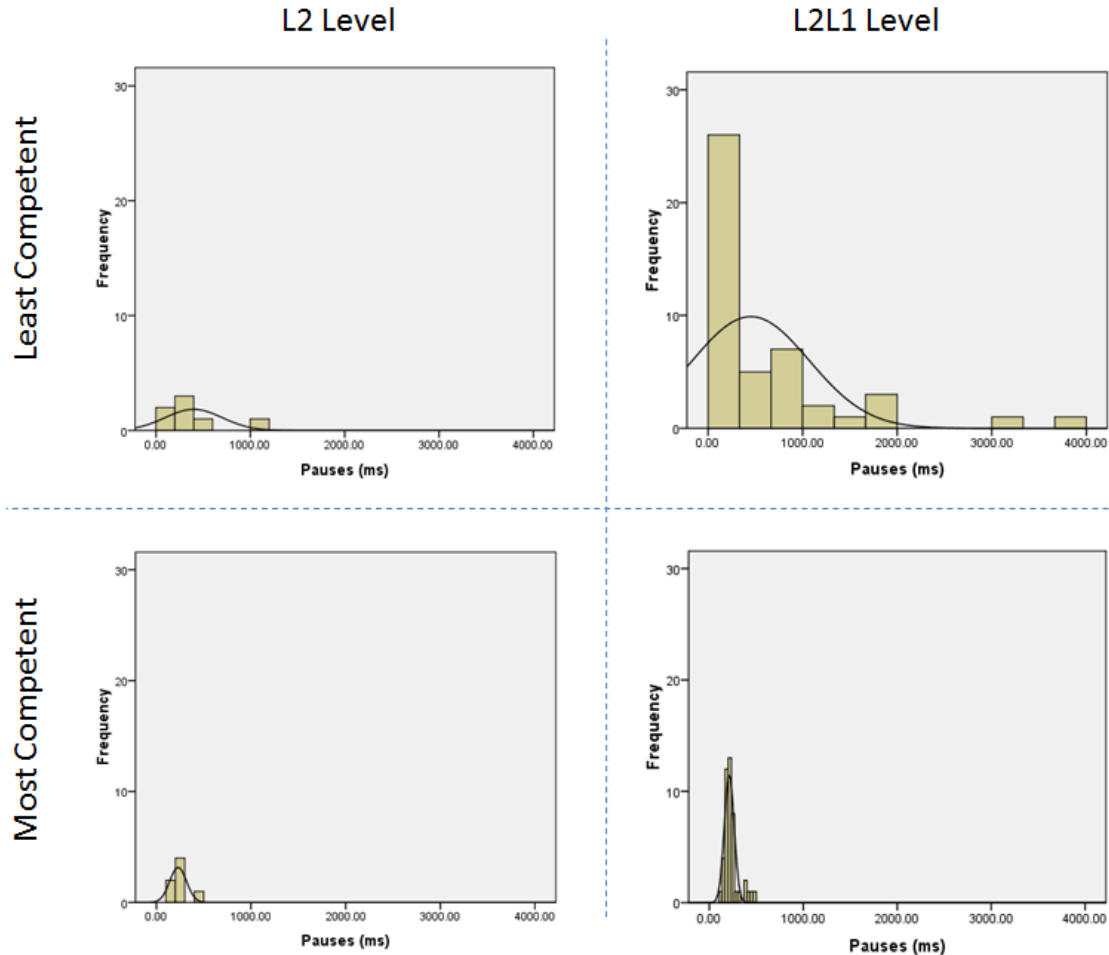


Figure 8.3 An example of pause distribution for Set III

Figure 8.3 shows the pause distribution for sentence 24, 'tanglerequirenaughtidleredictobtusenuptialepisode.' The scale here is different to those in Figure 8.1 and 8.2. With Set III, both L2 and L2L1 levels are still showing strong clustered pauses at the very low pause values for MC. However, there are differences in pause distributions for LC for both L2 and L2L1 Level. At L2 Level, the pauses are clustered at the low pause values, whereas at L2L1 level, the pauses are more widely distributed. This might explain why the correlations have changed so much, from being the lowest values with Method A and the highest values with Method B. What is particularly interesting about this group is that the distribution between L2 Level and L2L1 Level varies a lot for LC, but not for MC.

The graphs (Figures 8.1 to 8.3) show that changes in pause distributions are clearly seen for LC, but less so for MC. Finer detail analysis is needed at sentence level to examine why this is, and how LC copies the sentences.

The graphs of pause values were examined to observe the long pauses that are indicative of chunking. The chunks with long pauses were examined to see these long pauses were at L2 or L1 levels, in order to find the occurrence of syllable chunking (L1 pauses). It was found that Set I shows clear chunking mostly at L2 word level, Set II has chunking at both L2 (word) and L1 (syllable) position, while Set III shows clear chunking mostly at L1 (syllable) position. In order to further demonstrate these findings, three examples of the same sentence as that used in Figures 8.1–8.3 are investigated in more detail.

For Set I:

It was predicted that in the sentence **'itwouldbegoodformotherthoughithinkitcanbebetter'** from Set I, the omission of spaces would make it difficult for LC to distinguish the words, even though the words chosen are of high frequency, but that MC would not have a problem copying this sentence. Figure 8.4 shows a sample sentence with the pauses on the y-axis and the successive letters and words of the sentence on the x-axis. Each word is separated by a line on the x-axis, with the solid-line arrows pointing towards the long pauses that occur at L2 (word) level for LC

Results show that despite the predictions, LC is able to chunk at L2 (word) level, but takes a longer time to do so (as shown by the arrows). They also indicate that LC is able to distinguish the high frequency words. Some evidence of chunking that involves group of words is also seen with 'begood' and 'formother', though with 'thoughithink', the long pauses might have resulted from confusion in distinguishing the word. This confusion which might have been caused by the lack of familiarity of words or the inability to recognise words has resulted in the addition of letters to the words, for example 'thoughtitthink' or 'thoughitthink'. For some, the inability to accurately recognise words might also result in the missing of some words or letters. As in this case with letters being added, it can be seen by the pauses that are present for LC, but absent for MC. MC, however, produces short pauses of below 500ms throughout the sentence, and pause levels are mostly flat during the beginning of the copying. This might suggest the occurrence of parallel processing – the ability to copy and read the stimuli occurs simultaneously – because of the high

frequency of the words copied. Therefore, it can be concluded that when high frequency words are used, LC is able to distinguish those words, even when the spaces are omitted.

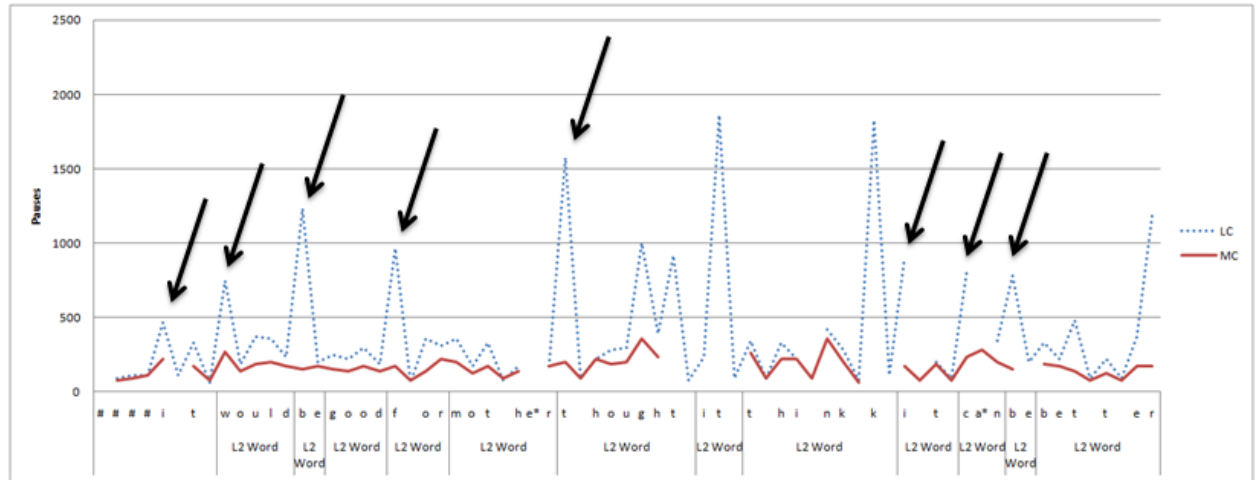


Figure 8.4 Example of Group I (sentence 19), where chunking can be mostly seen at L2 (word) level

For group II:

Given the sentence ‘**all actions that speak glitters louder is than not words gold**’ from group II, it was predicted that the words being clearly distinct through the spaces between them would help participants to comprehend the sentence. However, the jumbling of words, taken from two different proverbs, should make it difficult for LC to comprehend; it might look like a meaningless list of words. Hence, longer pauses might be seen at word level, associated with the attempt to comprehend the sentence meaning. Long pauses are predicted to occur within a sentence, when the words are considered to be long (more than 5 letters).

Figure 8.5 is a sample from Group II. Chunking can be seen at both L2 (word) and L1 (letter) level, especially for LC. The solid arrows show some examples of L2 (word) chunks, and the dashed arrows show some examples of L1 (letter or syllable) chunks. Although the specific point where L1 syllable chunking will occur cannot be identified, it is clear that it is taking place. It is highly likely that syllable chunking will occur, especially when participants are not familiar with the words; this is why L1 syllable chunking is usually seen to occur in long words. As seen, MC produces relatively short pauses – lower than 500ms, but does not have a flat range of pauses. This indicates that they might be trying to process the meaning, and the jumbled proverbs are affecting this process. LC, on the other hand, has many long pauses occurring at L2 or L1 level. Hence, it can be

concluded that both MC and LC are trying to comprehend the sentence, but at different levels of competence.

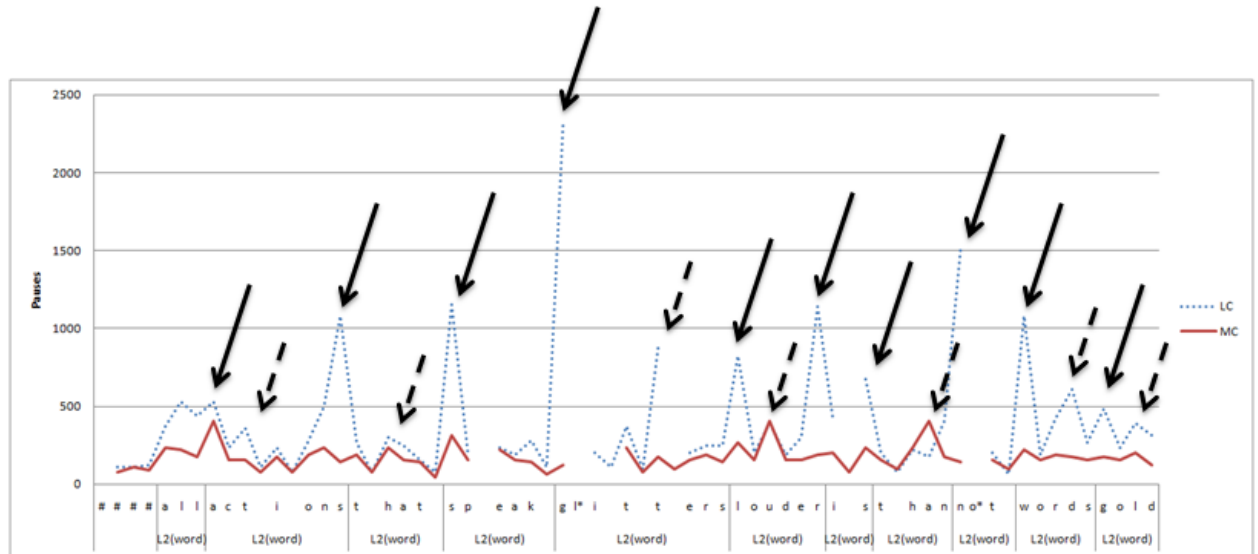


Figure 8.5 Example of Group II (sentence 21), where chunking can be seen at both L2 (word) and L1 (syllable) level

For group III:

In the sentence (24) **'tanglerequirenaughtidleredictobtusenuptialepisode'**, which contains low frequency words, no spaces, and each adjacent word containing high frequency bigrams at their end-letters, there was predicted to be an increased difficulty for both LC and MC in distinguishing the words. Looking at figure 8.6 proves these predictions. LC is shown to be carrying out lots of syllable chunking, with many long pauses occurring within words, and MC has some relatively flat lines which do not differentiate between L2s and L1s, suggesting that MC may or may not be able to recognise the words. However, MC benefits from having a greater familiarity of language that enables them to process copying with parallel processing; as they are capable of reading the stimulus while in the process of copying, a flat range of pauses is produced. However, when MC was copying the word 'naught', they might have realised that they were not able to distinguish the word. This state might have caused some disruption, hence the relatively long pause (although still under 500ms), and the possibility of being lost in terms of *copying cue* (i.e. forgetting where they last stopped reading), and hence missing copying the rest of the word 'idler'. However, MC caught up the copying cue within the word 'edict' and continued. LC, on the other hand, has more limited vocabulary knowledge and faces more difficulty in distinguishing the

words, hence producing lots of small chunks consisting of groups of letters, or syllables. This phenomenon, however, cannot be generalised; the processing will depend on the words and sentence type.

The weak correlations obtained from Method A (as in Table 8.2) may be because the method only measures the L2 pauses, and it has been seen that many long pauses occur within a word (L1 pauses). As Method B picks up all these L1 long pauses, it provides the strongest correlation for this particular sentence.

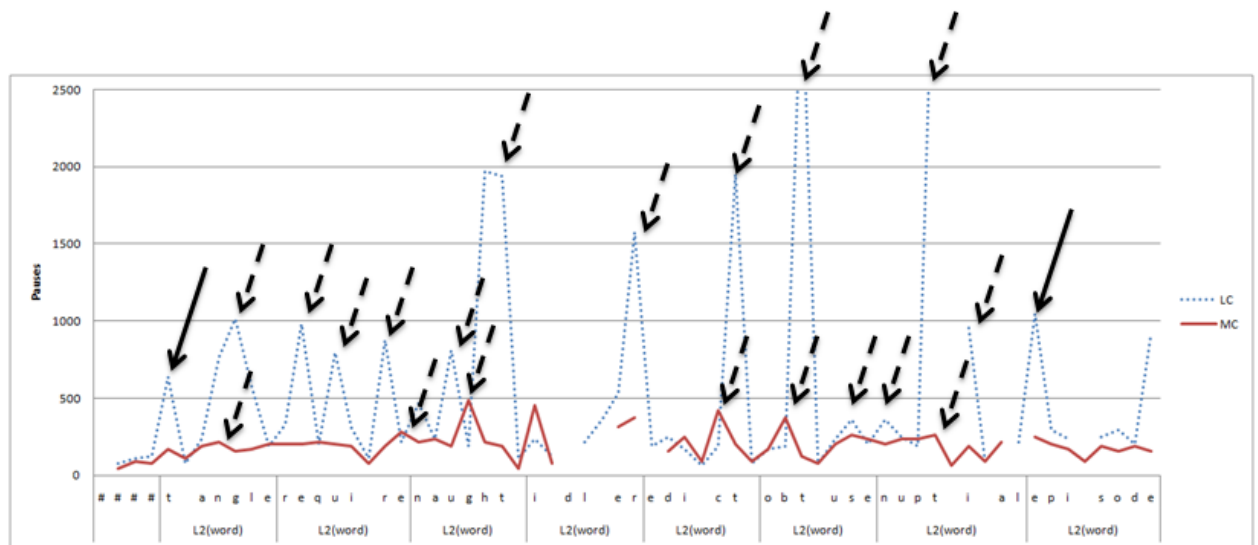


Figure 8.6 Example of Group III (sentence 24), where chunking can be seen mostly at L1 (syllable) level

### 8.2.2 Discussion for Experiment 1 Revisited

In conclusion to the revisiting of Experiment, it has been seen that at a general level, as in Table 8.1, Method B is not very promising, but when looking at a detailed analysis, it was discovered to be a potentially effective approach, with the right kind of sentence stimuli. It has been shown that using Method B, and hence involving L1 pauses, was able to show the occurrence of syllable effects and to distinguish the competent and the less competent.

The processing of syllables was found to be important, as this contributes to the processing time. Syllable process in shows that some participants may not be familiar with the words, or that they are familiar with high frequency syllables. Neglecting the L1 values, as in the case of L2Median measure, is not a good approach, because these data (L1 values) may contain important

information that relates to language processing. All the relevant data should be taken into account in order to assess a test taker.

This raises the question, however, of whether it is necessary to involve L1 at all. This may depend on the sentence stimulus. Method A might not be appropriate for certain types of stimuli as it does not examine syllable level chunking, but may be the most appropriate method for assessing word level processing. On the other hand, the theoretical model of copying (Chapter 4) supports the importance of L1 (letter or syllable) pauses, because every process consumes some amount of processing time and this may also involve syllable chunking.

Parallel processing was also observed; participants who are competent in the language are able to exploit their free processing resources to carry out parallel processing. For instance, they might be able to read the sentence stimulus at the same time as copying it down, and even move on to read the next stimulus whilst writing the current one. This phenomenon causes the absence of long pauses.

When the results of copying these sentences were compared with the baseline of copying alphabets (which is not meaningful in terms of language), it was clearly shown that the pauses from the baseline are not as flat or smooth as those found in the sentence-copying. Clear chunking is shown for groups of letters when copying the alphabet letters. This suggests that when carrying out parallel processing, MC is processing language with which they are greatly familiar. Even though the letters of the alphabet are familiar, copying the alphabet is not something that the MC does in everyday life. Comparing the alphabet writing baseline with the name writing, it was found that name writing has smoother and shorter pauses. This suggests that, unsurprisingly, the level of parallel processing that occurs during the sentence-copying is dependent on the participant's competence in the language.

To summarise the findings, the introduction of Method B, which incorporates L1 pauses, demonstrates the effects of syllable chunking, by measuring the long L1 (syllable level) pauses. Method A, which makes use of the Median of L2s, does not include L1 pauses, so with certain sentence types, it was not able to fully assess the participants' nature of language processing. Method B, on the other hand, was able to do so, particularly for the less competent participants, who did not have extended vocabulary knowledge. Method B did also evaluate competent

participants, however, especially when the sentence conditions (e.g. sentence numbers 24 and 25) rendered processing more difficult. Overall, it can be concluded that both measures are able to differentiate levels of competence, but in some cases this may depend on the type of sentences used.

### **8.3 Experiment 2 Revisited: Comparing Methods A & B in the First Trial**

Next, Experiment 2 is revisited, and the results of Methods A and B are again compared. Although there were issues with the effects of sentence order presentation in Experiment 2, the investigation was furthered by focusing only on the first trial of each participant to compare with the results obtained from Method A. Method B is hypothesised to be a better measure than Method A, and it was expected that the results from revisiting Experiment 2 will be able to prove this.

#### **8.3.1 Results**

In Experiment 2, the findings indicated that the sentence-copying might have been affected by the copying strategy issue, which largely depends on the order of the sentences and their jumbling conditions. This copying strategy employed by participants prevented further analysis, but an attempt was nevertheless made to analyse the data (in Chapter 6), by considering only the first trial performed by the participants, in which copying is free of any influence from sentence ordering effects. A similar approach was performed with Method B and the findings are summarised in Figure 8.7 below.

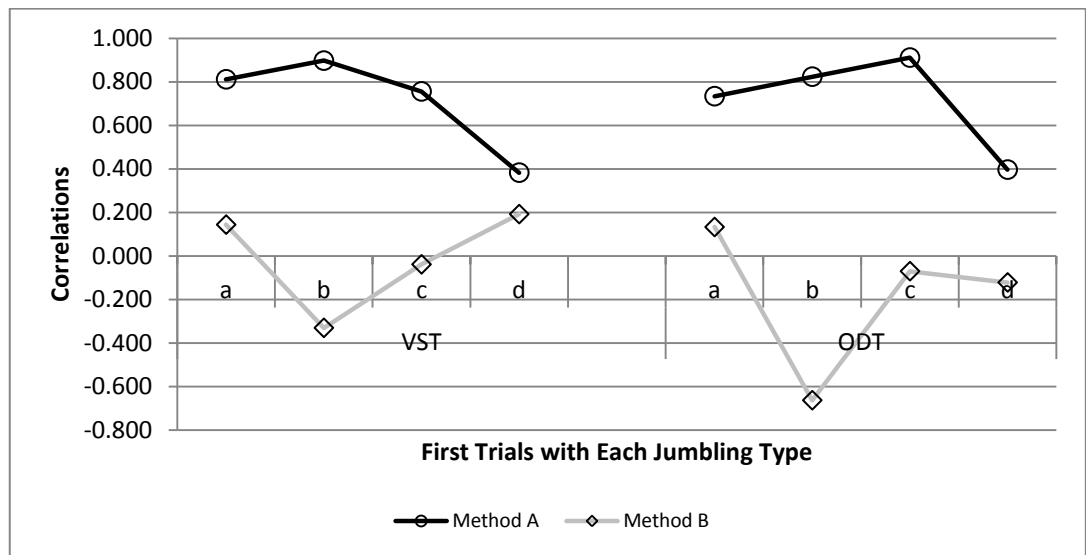


Figure 8.7 The first trial results of Methods A and B for Experiment 2 data

The correlation values are shown on the y-axis and the types of jumbling are shown on the x-axis, represented by the letters 'a' (*no jumbling*), 'b' (*words jumbled*), 'c' (*letters jumbled*), and 'd' (*both jumbled*). Twenty participants were divided among the four jumbling types, each group receiving a different jumbling type in their first trial. Thus, each jumbling type above consists of data from 5 participants. The darker lines with circle markers represent results for Method A, which involve only L2 (word) level pauses. The lighter line graphs with diamond markers represent results for Method B, which involves both L2 and L1 pauses. The left-hand graphs show correlations with VST and the right hand show correlations with ODT. Method A data here is data from after normalisation was performed; this was considered an improved version of the raw data.

The graphs clearly demonstrate the different interactions between Methods A and B. Applying Method B to this set of data does not improve the correlations; the Q3L2L1 is not as good as L2Medians. The inclusion of L1 in the calculation seems to have a different impact here than in the findings in Experiment 1. This is discussed further in the next section.

### 8.3.2 Discussion for Experiment 2 Revisited

Applying Method B with the data of Experiment 2 does not produce improved results over Method A. Involving L1 (letter or syllable) pauses seems to have affected the correlation values, from being significant with Method A to being not significant with Method B. These results may also be affected by the limited number of participants per sentence type group. There were 20



participants altogether, but only 5 participants did each of the jumbling types shown in Figure 8.7. Having only 5 participants in each group do not give a good range for picking out the most and the least competent. In the group that starts with jumbling type 'd' (*both jumbled*), the range of competence level is particularly close together. These data are therefore not considered suitable for performing the detailed analysis that was done with Experiment 1 revisited. It was concluded that in the case of Experiment 2, where there were effects of copying strategy and a limited number of participants for each jumbling condition, Method B could not be executed any further.

Comparing the stimuli used in Experiment 2 to the ones used in Experiment 1, those in Experiment 2 explores 4 different kinds of jumbling, which were based on findings of Experiment 1. Of the four jumbling types, *no jumbling* may belong to sentence Set I, *words jumbled* may belong to sentence Sets I or II, *letters jumbled* may belong to sentence Sets II or III, and *both jumbled* may belong to sentence Set III. These sentence Sets may act as a guideline to building up sentence stimuli for copying purposes in order to assess the cognitive processes of language processing.

## 8.4 Overall Discussion & Conclusion

This study set out to determine whether Method B is a better measure than Method A, which previous results suggested.

The results of this investigation show that Method B may not be as effective as Method A at a general level, but, a more detailed analysis (e.g. Section 8.2.1.2) was able to reveal its potential. However, Method B was found to be unsuccessful when implemented with Experiment 2 data, maybe because of the limitations of the number of participants recruited and the copying activities that were affected by the sentence order presentation. The opportunity to improve the Experiment 2 design, may allow further finer analysis may be able to be executed, and this may allow Method B to perform at its best. At this stage, this is still unclear, but it is certainly an issue which offers further investigation. The potential of implementing Method B, which lies in the fact that it shows meaningful findings at a detailed level, was clear from the positive findings of the comparison between the two methods with the data from Experiment 1.

Method B was able to show good results for certain sentence stimuli. Introducing L1 pauses has highlighted that participants tend to apply their syllable knowledge, and hence syllable chunking, when they face words which are long or unfamiliar. When developing a method to assess the cognitive processes, it is be important to take all levels of processing into account which could demonstrate an individual's nature of language processing.

Summing up, it is the researcher's opinion that there are still many open issues which need further investigation in the implementation of Method B, Q3L2L1. It is also possible that Methods A and B could be combined in order to provide an effective measure.

## CHAPTER 9 Discussion & Conclusion

### 9.1 Introduction

The overarching purpose of this research was to explore the use of pause analysis and sentence-copying for measuring the cognitive processes of language processing, which leads to the aim of this thesis:

**to develop a novel method for assessing components of cognitive processes in language processing by analysing the pauses that occurs during the activity of sentence-copying**

and that this is within the context of free handwriting and English as a second language.

A number of questions were posed in the introduction to this thesis (Chapter 1) with regard to achieving this aim, and a series of experiments were then conducted in order to answer them.

The questions posed are:

- (1) Why copying?
- (2) What kind of sentence stimuli should be copied?
- (3) What pause levels can determine language processes?
- (4) How should these pauses be calculated?
- (5) How should pauses be normalised to compensate for individual participants' differences?
- (6) How can language processes be measured?

These questions are built around the objectives of this thesis, which are to explore: (1) an effective method of copying, (2) the kinds of sentences suitable for copying, and (3) the possible measure of cognitive processes of language processing.

In order to try and understand the underlying cognitive processes involved when copying sentences, a Model of Copying (MoC) was developed. This MoC was also used for sentence design and to make predictions as to the reliability of the sentences as a tool for measuring language processes.

This final chapter draws together the findings of the present investigation in developing this novel method of measuring the cognitive processes of language processing. The chapter begins with the

discussion of the findings from all the experiments, based on the questions placed in the introduction chapter. The limitations, implications and future research contribution of each question are then discussed. The chapter concludes with a general statement stating the discoveries of this thesis.

## **9.2 Discussions of Findings, Limitations & Implications on the Future Work**

### **9.2.1 What is the Best Type of Copying?**

This study has shown that (1) Immediate Copying (IC) is a better choice than Initial Reading (IR), (2) there is a need to control the copying strategy that participants employ during the activity of copying, and (3) copying letters into grids of boxes is acceptable. These findings will be discussed one at a time, and the limitations of each will be examined, with suggestions for future research where appropriate.

Firstly, the comparison between IC and IR as copying techniques, in the pilot study, showed IC to be a better choice. IC provides better sets of data for measuring an individual's nature of language processing at that instant of time than IR, which allows some time for all levels of competence to familiarise themselves with the stimuli, giving them the opportunity to perform better and therefore providing a less accurate assessment of language processing. The findings showed that IC was able to provide stronger and more consistent correlations across all sentence types in the pilot than IR. IC was therefore applied from the early stages of this research. IC continued to provide strong correlations in subsequent experiments (1, 2 and 3) and this research can therefore be reasonably confident that IC is a suitable method of copying in measuring the cognitive processes of language processing.

Secondly, a need was found to control the way participants carry out the activity of copying, in order to gain more accurate data that only shows processes related to language competence, and does not include unrelated processing, such as that related to handwriting. Experiment 2 (Chapter 6) produced a limitation in terms of copying strategy effects affected by the sentence order presentation: the first jumbling condition received influenced the participants' strategy in copying the subsequent jumbling conditions. Experiment 3 (Chapter 7) therefore controlled the activity further by introducing two processing conditions, known as task types – processing by meaning

and fast-copying – instead of asking all participants to copy as quickly and as accurately possible. The aim in introducing these task types was to make sure that the strategy implemented in copying a particular sentence was not affected by the strategy adopted for the previous sentence: that the process related solely to that particular sentence type.

Although the introduction of these task types was shown in Experiment 3 to be reasonably effective, there was an issue with the True or False questions that were designed to confirm the processing by meaning. These questions were asked at the end of each copying task, based on the words in the sentence, so the question could perhaps be correctly answered simply by having spotted one key word in the sentence. By the time they copy the second, third and fourth sentences, participants might realise that they do not have to understand the overall meaning of the sentence, but only understand the shallow meaning of each word, in order to correctly answer the True or False questions. There is therefore a need to redesign these True or False questions in future research work.

Overall, this research has not been fully successful in controlling the task type in Experiment 3, and this is a critical aspect that needs to be developed further.

Thirdly, this study found the activity of copying the letters into individual boxes to be a useful method, showing strong correlations obtained throughout the experiments. This approach was chosen for this study as it showed strong results for Cheng & Rojas-Anaya (2004, 2005, 2006 and 2007) who first established the approach. Although it may be awkward to copy each letter into an individual box, the effect is consistent for every letter, so it should not have an effect on the overall data.

As this thesis research is concerned with pause analysis, there is a great need to distinguish the different levels of pauses, such as at word or letter level, in the production of copying. Distinguishing these levels of pauses would not be easy if cursive writing were used. It is important to be able to distinguish pauses at letter level in order to use these pauses to measure language processing. Similar intentions have been seen in studies from Kandel et al (2006) as well as Álvarez and Cottrell (2005), where they introduce the activity of pen lifting after each letter is copied, in order to try and show syllable chunking. Álvarez and Cottrell claimed that this helped to obtain clear data at the beginning and end of each letter.

Given the opportunity for further experiment, this research may have investigated asking participants to copy and write on a line, but with the instruction to make spaces between each individual letter, or maybe to lift the pen after each individual letter (i.e. non-cursive writing). Such an approach may be faster and produce less noise in the data. However, the new measure discovered, (Q3L2L1) may be able to automatically detect long pauses occurring within a word, so using cursive writing could perhaps still be effective with this measure.

### 9.2.2 What Kind of Sentence Stimuli?

Throughout the experiments, the study has found that there are some sentence types which might be more suitable for measuring the cognitive processes of language processing. The sentences explored vary in terms of their manipulations and also the type of sentence itself. ‘Manipulations’ may involve the jumbling of words or letters, the omission of spaces between words, the use of bigrams to combine words, and the use of words at different frequency levels. ‘Sentence type’ may mean an ordinary sentence, a simple sentence, a technical sentence, a proverb or a garden path sentence. The pilot and Experiment 1 explored a different range of these sentence types. However, the sentence type was standardised from Experiment 2 onwards, focusing on academic sentences, taken from journals.

First, strong correlations were found when using an ordinary sentence (*no jumbling*). This is apparent in all the experiments, including the pilot, and particularly so in Experiment 1, wherein the study established a suitable sentence stimuli. It was also shown that garden path sentences were processed in a similar way to ordinary sentences, due to the absence of long pauses at the ambiguity point within the sentence. The strong correlations obtained with this sentence type may indicate that participants were able to comprehend the sentence. This impression is supported further by the fact the garden path sentences still retain the correct grammatical structure and spellings, which helps to make the copying smooth.

Secondly, Experiment 1 showed that using jumbling conditions, such as the jumbling of words and letters, may also produce a suitable sentence stimulus. The findings of the pilot and Experiment 1 shaped the rest of the experiments to focus on *no jumbling*, *words jumbled*, *letters jumbled* and *both jumbled*. *No jumbling* preserves the grammatical structure and word spellings; *words jumbled* disrupts the grammar but preserves the spellings; *letters jumbled* disrupts the spellings but preserves the grammar; finally, *both jumbled* disrupts both the grammar and spellings.

The academic sentence used is aimed at making the stimuli design consistent across the copying tasks. Focusing only on one type of sentence allows the level of difficulty in terms of word frequencies to be controlled, for example by having a balance number of high- and low-frequency words. Participants' levels of familiarity with the technical terms used can then be assessed.

Finally, the choice of sentences and manipulations influences how an individual performs chunking. This is an important point to take into consideration in designing the stimuli because it affects the output. The intended output must tally with what is being measured. For example, if one wants to assess the comprehension level of reading a sentence (supported by good grammar structure), then jumbling letters would be irrelevant because this relates to word recognition and does not assess grammar comprehension. To conclude, in order to produce effective sentence stimuli, the difficulty of copying can be increased by carrying out different kinds of manipulations, but serious consideration must be made as to how this might affect the copying process.

### 9.2.3 Can Components of Language be Distinguished?

The correlations conducted against two independent tests, ODT and VST, showed that the sentence stimuli used were able to assess two language components: grammar and vocabulary. ODT, used from Experiment 1 onwards, assesses grammar knowledge, and VST, used from Experiment 2 onwards, assesses vocabulary size. Some sentences produced strong correlations with ODT and some with VST. This indicates that in some circumstances and in some stimuli, one language component or the other is being tested. The method in this thesis only examines processing at sentence level, but is able to measure both grammar and vocabulary knowledge, even with components that are not grammatically structured (i.e. the words are jumbled).

### 9.2.4 What Kind of Measure is Suitable?

The study found two measures that are suitable for calculating the pauses in order to get a single measure of language processing. These two measures are L2Median (the median of L2 pauses) and Q3L2L1 (the third quartile of the combination of L2 and L1 pauses). The decision of which measure is best to apply depends on the sentence stimuli used. L2Median focuses only on L2 pauses, while Q3L2L1 examines both L2 and L1 pauses. The third quartile of Q3L2L1 is able to measure only the long pauses, which are considered to be more meaningful than the short pauses. Short pauses usually represent the process of the writing of letters, whilst the long pauses

may represent the results of chunking at word group, word, syllable, or letter level. L2Median is not able to detect processes occurring at syllable level. As a conclusion, Q3L2L1 provides richer data than L2Median because it does not disregard long pauses which may occur at syllable level.

In the attempt to revisit previous experiments, Q3L2L1 was able to provide strong correlations for some of the sentence types, but these correlations were not as good as L2Median for some conditions. The revisit showed that applying both measures (L2Median and Q3L2L1) reveals three types of processes used by participants in copying the sentences. These were identified through a detailed analysis of the pause distributions of L2s only and of the combination of L2s and L1s, where they are: (1) clear word level processing, producing many long L2 pauses (2) both word and syllable level processing, producing long L2 and L1 pauses, and (3) clear syllable level processing, producing many long L1 pauses. This finding provides a guide to assigning a particular measure to a particular sentence condition. Chapter 8 identified sets of sentence types, grouped by the way in which correlation values change between applying L2Median and Q3L2L1. This shows that, for example, L2Median may suit sentences where words can be easily identified such as Set I, while Q3L2L1 may suit sentences where it is difficult to distinguish words such as those in Set II and III.

In terms of the practicality of using the measure, Q3L2L1 provides a much simpler approach than L2Median, and does not require complex data extraction. The complex data extraction required for L2Median means that one must separate the raw data by grouping the pauses into different pause levels: L2 for word level pauses, L1 for letter level pauses and L0 for stroke level pauses. Performing these data extractions takes a lot of time and is very tedious. It is also difficult to match the pause of each stroke to individual characters, because different participants have different ways of producing a letter. For example, some might use two strokes in producing the letter 'd', but others may use only one stroke. Further, carrying out the data extraction may be a problem if participants make mistakes, for example missing out letters in words, which would create a gap in the data. These must be manually corrected by the researcher. In Q3L2L1, however, the pauses can be simply calculated without the need to extract the data, so it is not necessary to code everything separately because of small mistakes such as adding or omitting a letter.

As a conclusion, even though both measures were shown to be useful, further investigation may be needed to confirm the effectiveness of Q3L2L1. These findings make the development of a



novel method for assessing components of language processing easier, because the newly-found measure, Q3L2L1, is useful in (1) excluding irrelevant pauses and (2) eliminating the hard work involved in distinguishing the pause levels manually or through a series of coding, instead allowing the inclusion of all pause data. In addition, the new measure discovered makes the approach in this thesis more feasible to apply.

### 9.2.5 Does Data Need to be Normalised?

With the newly-found measure, the study has found that there is no need to normalise because of the combination of L2 and L1 pauses in Q3L2L1. As discussed in the previous section, there exist for certain sentence stimuli long pauses within a word which are only transparent when using Q3L2L1. With L2Median, these pauses, which show important syllable processing, occurring within a word, are not shown. Therefore, Q3L2L1 provides a more straightforward measure than L2Median.

Chapter 5 of this research considered examining *pause difference* and *pause quotient*, as part of an attempt to compensate the data as a means to normalise, by subtracting (*pause difference*) or dividing (*pause quotient*) the pauses using L2 and L1 pauses. Baseline copying tasks were also designed, which included name writing and alphabet writing in upper case and lower case, as a means to normalise data. However, observations showed that using baselines produced similar results to using L1s pauses taken from the sentence copying, so it was decided to use L1 pauses of the sentence stimuli to perform the normalisation, as this is more straightforward.

To conclude, the more recent measure examined, Q3L2L1, seems to be the most efficient in many ways, including in its ability to indirectly normalise without the need to perform a specific measure in order to compensate the pauses.

### 9.2.6 The Model of Copying

In Chapter 4, this study produced a model that assists with making predictions in order to design sentence stimuli, and allows the underlying cognitive processes of copying to be understood. The model was developed in order to help understand how individuals perform the copying activity. Although copying may seem like a simple task, it is in fact a complex activity involving a number of processes that need to be understood. These processes may represent an important aspect

relating to the nature of processing in language. It is therefore important to understand how copying can differentiate the nature of language processing.

Using the Model of Copying, 7 possible processing routes are shown with which a participant might engage during the course of copying. The route the participant uses depends on the familiarity level of what is being copied. The routes taken also reflect different language processing levels. The pause lengths represent processes where short pauses demonstrate competence in the language while long pauses demonstrate lack of competence in the language. This model not only assists in explaining what is happening during the process of copying, but also points out where the weaknesses of a less competent participant might be present.

The model also shows the possibility of parallel processing: where several processes are occurring at the same time. This occurs because the greater familiarity of the sentences allows a participant to copy while reading. Parallel processing produces flat pauses, where it is difficult to distinguish the difference between L2 and L1 pauses.

The model is crude, in that it uses many assumptions. At present, the model is able to give a *processing load prediction* (PLP), a value that is predicted to represent the copying processes in order to make predictions on how somebody would perform the copying. The PLP values are predicted to be similar to the real pause duration. Currently, the model is useful for designing the sentence stimuli. This needs to be further validated; there is a need to find out the real processing time of each route, but doing this would require a significant amount of further research work. For current purposes, it is sufficient to base on assumptions that each process does require some amount of time; therefore a value is assigned to each process. Since the purpose of developing the model was to only to help understand what is going on cognitively, the model is only applied during sentence designing, to see if the sentence stimuli are able to differentiate the levels of competence. However, the model also opens up many opportunities for research in the area of copying and handwriting processes.

### 9.2.7 Critical Issues

The experiments reported in this thesis focused on exploring pause analysis to cognitively assess components of cognitive processes in language processing. Throughout the exploration, certain aspects have been focused on that the researcher believes are important to achieve the overall

aim. Therefore, this research has its own limitations. These limitations must be considered within the context of the aim of this study, and it is important that further research building upon this study only does so with the same target methodology. While this is the case, the limitations of this research work may also act as a future research question.

The research limitations are:

1. The study only considers participants with English as their second language and SL learners who use alphabetical letters in their FL. Therefore, the work is limited in terms of the nature of the SL, such as the limitation of word frequency and vocabulary knowledge. The use of alphabetical letters is aimed at controlling any effects of orthography differences in their writing, although it may be interesting to include other character writing in the future. This limitation is purposefully done, as this study is the first of its kind. Although the thesis has restricted the participants' choice as such, the nature of SL itself is very wide, as discussed in Section 2.4, which has already given this research a good range of difference in competence. Future research may involve participants with English as their first language, but careful consideration is still needed in the choice of participants to avoid too large a range of difference in competence.
2. The study involves an unusual way of writing single letters in boxes, to avoid cursive writing and thus easily distinguish pauses at word and letter level. Future research may involve investigating writing on straight lines, but this must still be non-cursively, in order to distinguish the pauses at each letter.
3. The sentences used consisted of roughly 5 to 25 words, with the shortest being a proverb and the longest being an ordinary sentence with technical words. Future research might find it useful to use sentences that are all of the same length.
4. This overall exploration involves only adults between the ages of 20 and 45. Future research could involve a wider age group, such as children and senior adults.
5. The study only involves two groups of populations: Malaysian and Spanish. This was because of the need to use participants who use the same script of writing (i.e. the Roman alphabet). It may be a worthwhile piece of research to explore other populations who use the same scripts, or who use different scripts. Such may be interesting as to see (1) the effects of

handwriting production differences onto the copying activities and (2) the effects of first language on how a participant performs the chunking and copying processes.

6. The research does not deal with participants with language problems, such as dyslexia. This may be an interesting avenue of research.
7. This study only explores two approaches of copying: Immediate Copying (IC) and Initial Reading (IR). There may be other approaches which are useful for exploration, such as tracing or dictation.
8. Simple statistical measures have been explored in this study: median, mean and the quartiles. The complexity of these measures was increased by carrying out normalisation. However, further exploration could be carried out into other potentially effective measures.
9. This study is limited to the two independent tests used – ODT and VST – which determine two language components – grammar and vocabulary. Other useful tests may be able to point out other kinds of language components, such as spelling or a specific component of grammar (for example examining verbs or noun usage, or even punctuation).
10. This study only involves the common scripts of writing (letters of the alphabet). There may be research possibilities in exploring other language and writing scripts such as the Chinese, Indian or Arabic scripts.

The overall results of this research indicate that the application of pause analysis has some potential for measuring the cognitive processes of language processing in ESLs. Not only that, but the method of pause analysis and the copying of sentences allows researchers to understand the low-level cognitive processes of writing, which include studies of the impacts of writing. Although this research is still at an initial stage, it suggests many interesting potential future studies which may benefit the Second Language Acquisition and Cognitive Sciences fields. Further work is definitely needed to establish this novel method, especially for the verification of the Theoretical Model of Copying and the exploration of the sentences used.

### 9.3 Overall Conclusion

This thesis sought to establish a novel method for measuring components of cognitive processes in language processing. It has established a suitable copying method (IC), suitable sentence and manipulation types, and two suitable measures for assessing language processing (L2Median and

Q3L2L1). It has established that there is genuine potential for measuring language processing using pause analysis.

The research closes with a number of open issues that need further investigation to fully establish the effectiveness of the method. It creates opportunities for future research to build on the findings and more conclusively establish the method, using a wider or more diverse participant group, or a wider range of sentence types.

This research has demonstrated the potential of this novel method. The full development of the method will inevitably involve further research.

## Appendices

### Baseline Writing for Chapter 4 and 5

1 - A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

2 - a b c d e f g h i j k l m n o p q r s t u v w x y z

3 - m P S z b F g I N Q r A x J o U c e H T w k V I D y

4 - zjbmt lzp bvjh dw qmyjd fkz pvkx vm hbln zdfm

5 - Please write your name in small letters only; 3 times with no spaces in between.

6 - Please write your name all in capital letters only; 3 times with no spaces in between.

### Baseline Writing for Chapter 6 and 7

1 - Please write your name in small letters only; 3 times with no spaces in between.

2 - Please write your name all in capital letters only; 3 times with no spaces in between.

## Sentence Stimuli for Chapter 4 and 5

- 7 - If there is appreciable interaction between the normal coordinates of the activated complex such a short life will give, instead of sharp energy levels, the fuzzy quantization observed in predissociation.
- 8 - Wiretap subjects can also falsely indicate the ending times for calls they make and receive and can inject false records of outgoing and incoming calls into pen register logs.
- 9 -  
Many mutations can destabilize native conformations and promote formation of aggregates with properties of amyloid fibrils
- 10 -  
Length scale is a fundamental quantity that dictates the type of forces governing physical phenomena
- 11 - the other side greener the grass is always on
- 12 - are than bigger eyes your belly
- 13 - This is because the human mind does not read every letter by itself, but the word as a whole
- 14 - The beaver colony was neither large nor small, having neither females and ten members, and like all beaver they worked very hard to dam the small river
- 15 - The horse raced past the barn fell.
- 16 - The man who hunts ducks out on weekends.
- 17 - The psychologist told the wife that he was having trouble with to leave.
- 18 - what have she done now that she should tell mother first
- 19 - it would be good for mother though I think it can be better
- 20 - all a good friend things in must need come is to a an friend end indeed
- 21 - all actions that speak glitters louder is than not words gold
- 22 - birds every it of dog is a has raining feather his cats flock day and together dogs
- 23 - bad beauty appearances news is are travels only deceptive fast skin-deep
- 24 - tanglers require naughty dildoes to obtain a euphoric episode
- 25 - the artist intimidated distraught honorable evildoer dramatized narrator

## Sentence Stimuli for Chapter 6

1A	It normally involves having access to a secure site on the internet where a graded series of lessons are available.
1B	the it internet normally where involves a having graded access series to of a lessons secure are site available on
1C	it nolmarly inlovves hanivg acsecs to a seruce stie on the innertet wehre a gredad sereis of lensoss are avaiballe
1D	the it innertet nolmarly wehre inlovves a hanivg gredad acsecs sereis to of a lensoss seruce are stie avaiballe on
2A	Expert-oriented views of development distance the researcher from reality create barriers which promote ignorance and perpetuate inappropriate models.
2B	create expert-oriented barriers views which of promote development ignorance distance and the perpetuate researcher inappropriate from models reality
2C	exrept-oriethed viwes of demelopvent dintasce the rehearcser form reatily craete barreirs whcih protome igronance and petperuate inapptopriare moleds
2D	craete exrept-oriethed barreirs viwes which of protome demelopvent igronance dintasce and the petperuate rehearcser inapptopriare form moleds reatily
3A	The taxi is absolutely integral to the daily life of any urban fabric but it also incites fear.
3B	of the any taxi urban is fabric absolutely but integral it to also the incites daily fear life
3C	the txai is abtolusely ingetral to the daliy lfie of any urabn farbic but it aslo intices faer
3D	of the any txai urabn is farbic abtolusely but ingetral it to aslo the intices daliy faer lfie
4A	It is a space that is activated via identificatory strategies that are always external to it.
4B	identificatory it strategies is that a are space always that external is to activated it via
4C	it is a spcae taht is actitaved via idencifitatory strageties taht are alyaws exnertal to it
4D	idencifitatory it strageties is taht a are spcae alyaws taht exnertal is to actitaved it via
5A	Organisational competence consists of grammatical and textual competencies while pragmatic competence consists of illocutionary and sociolinguistic competencies.
5B	while organisational pragmatic competence competence consists consists of of illocutionary grammatical and and sociolinguistic textual competencies competencies
5C	organinatiosal cotpemence cossints of gramcatimal and tetxual comcetenpies whlie pratmagic competemce cosnists of iltoculionary and sogiolincuistic cotpemencies



5D	while organizational pragmatic competence consists of of intercultural grammatical and sociolinguistic textual competencies
6A	Baby boomers are less satisfied with their overall health report more pain and cite more difficulty with routine activities than their predecessors.
6B	pain baby and boomers cite are more less difficulty satisfied with with routine their activities overall than health their report predecessors more
6C	bbay boorems are lses safistied wth thier ovelarl heatlh reropt mroe pian and ctie mroe dilficulty wthi rounite activitvies tahn thier precedessors
6D	pian bbay and boorems ctie are mroe lses dilficulty safistied wthi wthi rounite thier activitvies ovelarl tahn heatlh thier reropt precedessors mroe
7A	This engagement is especially important among predominantly young designers and developers in the wireless industry.
7B	young this designers engagement and is developers especially in important the among wireless predominantly industry
7C	tihs enmagegent is escepially imtorpant amnog prenomidantly yonug desingers and delevopers in the wileress intusdry
7D	yonug tihs desingers enmagegent and is delevopers escepially in imtorpant the amnog wileress prenomidantly intusdry
8A	The representation of dramatic narrative in pictorial rather than in visual experiential terms is evident in images.
8B	than the in representation visual of experiential dramatic terms narrative is in evident pictorial in rather images
8C	the retresenpation of dratamic natrarive in pirtocial rahter tahn in viusal expetienrial temrs is evinedt in imegas
8D	tahn the in retresenpation viusal of expetienrial dratamic temrs natrarive is in evinedt pirtocial in rahter imegas
9A	Comparing human and monkey gaze behavior with a computational saliency model revealed interspecies gaze correlations.
9B	a comparing computational human saliency and model monkey revealed gaze interspecies behavior gaze with correlations
9C	copmaring huamn and mokney gzae bevahior wthi a cotpumational sacienly moedl releaved inperstecies gzae colrerations
9D	a copmaring cotpumational huamn sacienly and moedl mokney releaved gzae inperstecies bevahior gzae wthi colrerations

10A	Resolving wicked problems whether they be simple or complex requires an approach that is participatory and adaptive.
10B	complex resolving requires wicked an problems approach whether that they is be participatory simple and or adaptive
10C	revolsing wikced promlebs whehter tehy be silpme or colpmex reruiques an apcroaph taht is patticiparory and adaptive
10D	colpmex revolsing reruiques wikced an promlebs apcroaph whehter taht tehy is be patticiparory silpme and or adaptive
11A	The struggle involved much ravaging and numerous small skirmishes these are hinted at in early Welsh poetry.
11B	skirmishes the these struggle are involved hinted much at ravaging in and early numerous Welsh small poetry
11C	the stluggre invovled mcuh ragaving and nuremous smlal skimrishes thsee are hitned at in ealry Weslh poerty
11D	skimrishes the thsee stluggre are invovled hitned mcuh at ragaving in and ealry nuremous Weslh smlal poerty
12A	It is because rationalised subjectivity has the potential to encounter a repressed dimension of itself that musical material can embody a reconfigured subjectivity.
12B	repressed it dimension is of because itself rationalized that subjectivity musical has material the can potential embody to a encounter reconfigured a subjectivity
12C	it is besauce ralionatised suvjectibity has the potetnial to entouncer a repesred disenmion of itlesf taht mucisal maretial can emdoby a regonfigured a subvectijity
12D	repesred it disenmion is of besauce itlesf ralionatised taht suvjectibity mucisal has maretial the can potential emdoby to a entouncer regonfigured a subvectijity

## Sentence Stimuli for Chapter 7

NO	SENTENCES	Questions
1a	the representation of dramatic narrative in pictorial rather than in visual experiential terms is evident in images	Is the sentence about a 'drama festival'? <input type="radio"/> True <input type="radio"/> False.....ANSWER
1b	than the in representation visual of experiential dramatic terms narrative is in evident pictorial in rather images	Is the sentence about a 'drama festival'? <input type="radio"/> True <input type="radio"/> False.....ANSWER
1c	the retresenpation of dratamic natrarive in pirtocial rahter tahn in viusal expetiennial temrs is evinedt in imegas	Is the sentence about a 'drama festival'? <input type="radio"/> True <input type="radio"/> False.....ANSWER
1d	tahn the in retresenpation viusal of expetiennial dratamic temrs natrarive is in evinedt pirtocial in rahter imegas	Is the sentence about a 'drama festival'? <input type="radio"/> True <input type="radio"/> False.....ANSWER
2a	comparing human and monkey gaze behavior with a computational saliency model revealed interspecies gaze correlations	Is the sentence about an 'experimental study'? <input type="radio"/> True.....ANSWER <input type="radio"/> False
2b	a comparing computational human saliency and model monkey revealed gaze interspecies behavior gaze with correlations	Is the sentence about an 'experimental study'? <input type="radio"/> True.....ANSWER <input type="radio"/> False
2c	copmaring huamn and mokney gzae bevahior wtih a cotpumational sacienly moedl releaved inperstecies gzae colrerations	Is the sentence about an 'experimental study'? <input type="radio"/> True.....ANSWER <input type="radio"/> False
2d	a copmaring cotpumational huamn sacienly and moedl mokney releaved gzae inperstecies bevahior gzae with colrerations	Is the sentence about an 'experimental study'? <input type="radio"/> True.....ANSWER <input type="radio"/> False
3a	it is a space that is activated via identificatory strategies that are always external to it	Is the sentence about a 'research by NASA'? <input type="radio"/> True <input type="radio"/> False.....ANSWER
3b	identificatory it strategies is that a are space always that external is to activated it via	Is the sentence about a 'research by NASA'? <input type="radio"/> True <input type="radio"/> False.....ANSWER
3c	it is a spcae taht is actitaved via idencifitatory strageties taht are alays exnertal to it	Is the sentence about a 'research by NASA'? <input type="radio"/> True <input type="radio"/> False.....ANSWER
3d	idencifitatory it strageties is taht a are spcae alays taht exnertal is to actitaved it via	Is the sentence about a 'research by NASA'? <input type="radio"/> True <input type="radio"/> False.....ANSWER

		<input type="radio"/> True <input type="radio"/> False..... <b>ANSWER</b>
<b>4a</b>	it normally involves having access to a secure site on the internet where a graded series of lessons are available	Is the sentence about a ' <b>protected webpage</b> '? <input type="radio"/> True..... <b>ANSWER</b> <input type="radio"/> False
<b>4b</b>	the it internet normally where involves a having graded access series to of a lessons secure are site available on	Is the sentence about a ' <b>protected webpage</b> '? <input type="radio"/> True..... <b>ANSWER</b> <input type="radio"/> False
<b>4c</b>	it nolmarly inlovves hanivg acsecs to a seruce stie on the innertet wrehe a gredad sereis of lensoss are avaialle	Is the sentence about a ' <b>protected webpage</b> '? <input type="radio"/> True..... <b>ANSWER</b> <input type="radio"/> False
<b>4d</b>	the it innertet nolmarly wrehe inlovves a hanivg gredad acsecs sereis to of a lensoss seruce are stie avaialle on	Is the sentence about a ' <b>protected webpage</b> '? <input type="radio"/> True..... <b>ANSWER</b> <input type="radio"/> False
<b>5a</b>	this engagement is especially important among predominantly young designers and developers in the wireless industry	Is the sentence about a ' <b>wedding preparation</b> '? <input type="radio"/> True <input type="radio"/> False..... <b>ANSWER</b>
<b>5b</b>	young this designers engagement and is developers especially in important the among wireless predominantly industry	Is the sentence about a ' <b>wedding preparation</b> '? <input type="radio"/> True <input type="radio"/> False..... <b>ANSWER</b>
<b>5c</b>	tihs enmagegent is escepiially imtorpant amnog prenomidantly yonug desingers and delevelopers in the wileress intusdry	Is the sentence about a ' <b>wedding preparation</b> '? <input type="radio"/> True <input type="radio"/> False..... <b>ANSWER</b>
<b>5d</b>	yonug tihs desingers enmagegent and is delevelopers escepiially in imtorpant the amnog wileress prenomidantly intusdry	Is the sentence about a ' <b>wedding preparation</b> '? <input type="radio"/> True <input type="radio"/> False..... <b>ANSWER</b>
<b>6a</b>	organisational competence consists of grammatical and textual competencies while pragmatic competence consists of illocutionary and sociolinguistic competencies	Is the sentence about ' <b>language</b> '? <input type="radio"/> True..... <b>ANSWER</b> <input type="radio"/> False
<b>6b</b>	while organisational pragmatic competence competence consists consists of of illocutionary grammatical and and sociolinguistic textual competencies competencies	Is the sentence about ' <b>language</b> '? <input type="radio"/> True..... <b>ANSWER</b> <input type="radio"/> False
<b>6c</b>	organinatiosal cotpemence cossints of gramcatimal and tetxual comctenpies whlie pratmagic competemce cosnists of iltoculinary and sogiolincuistic cotpemencies	Is the sentence about ' <b>language</b> '? <input type="radio"/> True..... <b>ANSWER</b> <input type="radio"/> False
<b>6d</b>	whlie organinatiosal pratmagic cotpemence competemce cossints cosnists of of iltoculinary gramcatimal and and sogiolincuistic tetxual cotpemencies comctenpies	Is the sentence about ' <b>language</b> '? <input type="radio"/> True..... <b>ANSWER</b> <input type="radio"/> False

<b>7a</b>	baby boomers are less satisfied with their overall health report more pain and cite more difficulty with routine activities than their predecessors	Is the sentence about ' <b>children's development</b> '? <input type="radio"/> True <input type="radio"/> False..... <b>ANSWER</b>
<b>7b</b>	pain baby and boomers cite are more less difficulty satisfied with with routine their activities overall than health their report predecessors more	Is the sentence about ' <b>children's development</b> '? <input type="radio"/> True <input type="radio"/> False..... <b>ANSWER</b>
<b>7c</b>	bbay boorems are lses safistied wtih thier ovelarl heath reropt mroe pian and ctie mroe dilficulty wtih rounite actitivies tahn thier precedessors	Is the sentence about ' <b>children's development</b> '? <input type="radio"/> True <input type="radio"/> False..... <b>ANSWER</b>
<b>7d</b>	pian bbay and boorems ctie are mroe lses dilficulty safistied wtih wtih rounite thier actitivies ovelarl tahn heatlh thier reropt precedessors mroe	Is the sentence about ' <b>children's development</b> '? <input type="radio"/> True <input type="radio"/> False..... <b>ANSWER</b>
<b>8a</b>	expert-oriented views of development distance the researcher from reality create barriers which promote ignorance and perpetuate inappropriate models	Is the sentence about ' <b>marketing</b> '? <input type="radio"/> True <input type="radio"/> False..... <b>ANSWER</b>
<b>8b</b>	create expert-oriented barriers views which of promote development ignorance distance and the perpetuate researcher inappropriate from models reality	Is the sentence about ' <b>marketing</b> '? <input type="radio"/> True <input type="radio"/> False..... <b>ANSWER</b>
<b>8c</b>	exrept-orientned viwesof demelopvent dintasce the rehearcser form reatily craete barreirs which protome igronance and petperuate inapptopriare moleds	Is the sentence about ' <b>marketing</b> '? <input type="radio"/> True <input type="radio"/> False..... <b>ANSWER</b>
<b>8d</b>	craete exrept-orientned barreirs viwes whcih of protome demelopvent igronance dintasce and the petperuate rehearcser inapptopriare form moleds reatily	Is the sentence about ' <b>marketing</b> '? <input type="radio"/> True <input type="radio"/> False..... <b>ANSWER</b>

## References

- Ahlsen, E. and Stromqvist, S. (1999). *ScriptLog: A tool for logging the writing process and its possible diagnostic use. Paper presented at the 1998 ISAAC Research Symposium, Dublin, Ireland*
- Akamatsu, N. (1999). The effects of first language orthographic features on word recognition processing in English as a second language. *Reading and Writing: An Interdisciplinary Journal*, 11(4), 381–403.
- Alamargot, D., & L. Chanquoy. (2001). *Through the models of writing*. Dordrecht: Kluwer.
- Alamargot, D., Chesnet, D., Dansac, C., & Ros, C. (2006). Eye and pen: A new device for studying reading during writing. *Behavior Research Methods, Instruments, & Computers*, 38, 287-299.
- Alderson, J.C. & Banerjee, J. (2001). Language testing and assessment (Part 1). *Language Teaching*, [Online].34(213-236), Available at: <http://eprints.lancs.ac.uk/1036/1/displayFulltext2.pdf> [Accessed August 2009]
- Alderson, J.C. & Banerjee, J. (2002). Language testing and assessment (Part 2). *Language Teaching*, [Online].3(213-236), Available at: <http://journals.cambridge.org> (Cambridge Journals Online) [Accessed August 2009]
- Alderson, J.C. (2000). *Assessing reading ability*. Cambridge: Cambridge University Press.
- Alexander, P.A., & Jetton, T.L. (2000). Learning from text: A multi-dimensional and developmental perspective. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research: Vol. III* (pp. 285-310). Mahwah, NJ: Erlbaum.
- Álvarez, C.J., Carreiras, M., de Vega, M. (2000). Syllable-frequency effect in visual word recognition: Evidence of sequential-type processing. *Psicológica* 21. 3: 341-374.
- Anderson, N. J. (1991). Individual differences in strategy use in second language reading and testing. *The modern language journal*, 75(4), 460-472.
- Anderson, R.C. and Freebody, P. (1979). Vocabulary knowledge. (ERIC Document Reproduction Service No. ED 177 480).
- Andersson, B., Dahl, J., Holmqvist, K., Holsanova, J., Johansson, V., Karlsson, H. (2006). Combining keystroke logging with eye tracking. In L. Van Waes, M. Leiten, & C. M. Neuwirth (Eds.), *Writing and digital media* (pp. 166-172). Amsterdam: Elsevier.
- Askov, E.N. (1975) Handwriting: copying versus tracing as the most effective type of practice. *Journal of*

Education Resources 69:96–98

- Baddeley, A.D. & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (vol.8, pp. 281-301). London: Academic Press.
- Baddeley, A.D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short term memory. *Journal of Verbal Learning and Verbal Behaviour*, 14, 575-589
- Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 28±61). Cambridge, UK: Cambridge University Press.
- Barbe, W.B., Lucas, V.H. & Wasylyk, Th.M. (1984). *Handwriting: Basic skills for effective communication*, Zaner-Bloser, Columbus, OH
- Barone, R. (2010) (work in progress)
- Birch, B. (2007). *English L2 reading: Getting to the bottom* (2<sup>nd</sup> ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Bonin, P. (1997). Produire des mots isolés oralement et par écrit. *Revue de Neuropsychologie*, 7, 29–70.
- Bonin, P., Fayol, M., & Gombert, J.E. (1997). Role of phonological and orthographic codes in picture naming and writing: An interference paradigm study. *Current Psychology of Cognition*, 16, 299–320.
- Bourdin, B., & Fayol, M. (1994). Is written language production more difficult than oral language production? A working memory approach. *International Journal of Psychology*, 29, 591–620.
- Bridwell, L. S., Sirc, G., & Brooke, R. (1985). Revising and composing: Case studies of student writers. In S. Freedman (Ed.), *The acquisition of written language: Revision and response* (pp. 172–194). Norwood, NJ: Ablex.
- Bridwell-Bowles, L. S., Johnson, P. and Brehe, S. (1987). Computers and composing: Case studies of experienced writers. In A. Matsuhashi (ed.) *Writing in Real Time: Modelling Production Processes* (pp. 81-107). New York: Longman.
- Brinkmann, E. (2004). Abschreiben üben. [Practising copying.]. *Grundschule Deutsch*, 2, 29.
- Butterworth, B. (1980). Some constraints on models of language production. In B. Butterworth (Ed.), *Language production: Vol. 1. Speech and talk*. New York: Academic Press.
- Carroll, J.B. (1965). The prediction of success in intensive foreign language training. In R.Glaser (Ed.), *Training, research, and education* (pp. 87-136). New York: Wiley.
- Chanquoy, L., Foulon, J.-N., & Fayol, M. (1990). Temporal management of short text writing by children

and adults. *Cahiers de Psychologie Cognitive*, 10, 513-540.

- Cheng, P. C. H., & Rojas-Anaya, H. (2005). Writing out a temporal signal of chunks: patterns of pauses reflect the induced structure of written number sequences. In B. G. Bara, L. Barsalou & M. Bucciarelli (Eds.), *Proceedings of the Twenty Seventh Annual Conference of the Cognitive Science Society* (pp. 424-429). Mahwah, NJ: Lawrence Erlbaum.
- Cheng, P. C. H., & Rojas-Anaya, H. (2006). A temporal signal reveals chunk structure in the writing of word phrases. In *Proceedings of the Twenty Eighth Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum.
- Cheng, P. C. H., & Rojas-Anaya, H. (2007). Measuring Mathematical Formula Writing Competence: An Application of Graphical Protocol Analysis. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the Twenty Ninth Annual Conference of the Cognitive Science Society* (pp. 869-874). Austin, TX: Cognitive Science Society.
- Cheng, P. C. H., & Rojas-Anaya, H. (2008). A Graphical Chunk Production Model: Evaluation Using Graphical Protocol Analysis with Artificial Sentences. In B. C. Love, K. McRae & V. M. Sloutsky (Eds.), *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society* (pp. 1972-1977). Austin, TX: Cognitive Science Society.
- Cheng, P. C.-H., McFadzean, J., & Copeland, L. (2001). Drawing out the temporal structure of induced perceptual chunks. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the Twenty Third Annual Conference of the Cognitive Science Society* (pp. 200-205). Mahwah, New Jersey: Lawrence Erlbaum.
- Cheng, P.C-H., & Rojas-Anaya, H. (2004). TRACE user guide (Unpublished Representational Systems Laboratory report).
- Christianson, K., Johnson, R. L., & Rayner, K. (2005). Letter transposition within and across morphemes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 1327-1339
- Cohen, A. D., & Upton, T. A. (2006). *Strategies in responding to new TOEFL reading tasks*. TOEFL Monograph Series (MS-33). Princeton, NJ: ETS.
- Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151-216). New York: Academic Press.
- Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel distributed-processing approaches. *Psychological Review*, 100, 589-608.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of



- visual word recognition and reading aloud. *Psychological Review*, 108(1), 204–256.
- Cook, V. (1992). Evidence for multicompetence. *Language Learning*, 42(4):557-91.
- Cowan, N. (1988) Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin* 104:163–91.
- Cowan, N. (1995) *Attention and memory: An integrated framework*. Oxford Psychology Series, No. 26. Oxford University Press.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87-114.
- De Beaugrande, R. (1984) *Text production: Toward a science of composition*. Norwood, NJ: Ablex. (p.160)
- DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing system*. Honolulu: University of Hawai'i.
- E. Askov, W. Otto, W. Askov (1970) A decade of research in handwriting: Progress and prospect. *Journal of Educational Research*, 64, pp. 100–111
- Ehri, L. C. (2000). Learning to read and learning to spell: Two sides of a coin. *Topics in Language Disorders* 20(3), 19-36.
- Ehri, L.C. (1994). Development of the ability to read words: Update. In R. Ruddell, M. Ruddell & H. Singer (Eds.), *Theoretical models and processes of reading*. (4th Edition, pp. 323–358). Newark, Del: International Reading Association.
- Ehri, L.C. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J.L. Metsala & E.C. Ehri (Eds.), *Word recognition in beginning literacy*. (pp. 3–40). Mahwah, NJ: Erlbaum.
- Ehri, L.C. (2005). Development of sight word reading: Phases and findings. In M.J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook*. (pp. 135–154). Oxford: Blackwell.
- Ellis, A.W. (1982). *Spelling and writing (and reading and speaking), Normality and pathology in cognitive functions*, Academic Press, London
- Ellis, A.W. (1988) Normal writing processes and peripheral acquired dysgraphias, *Language and Cognitive Processes*, 3, pp. 99–127
- Ehri, L. C., & Wilce, L. S. (1980). The influence of orthography on readers' conceptualization of the phonemic structure of words. *Applied Psycholinguistics*, 1(4), 371-385.

- Ellis, A.W., Young, A.W. (1988). Human cognitive neuropsychology. Erlbaum, London (1988)
- Fang, S.-P., Horng, R.-Y., & Tzeng, O. J. L. (1986). Consistency effects in the Chinese character and pseudo-character naming task. In H. S. R. Kao & R. Hoosain (Eds.), *Linguistics, psychology, and the Chinese character* (pp. 11–21). Hong Kong: University of Hong Kong, Center of Asian Studies.
- Fayol, M. (1998). From on-line management problems to strategies in written composition. In G. Rijlaarsdam & E. Esperet (Series Ed.) & G. Jefferey (Vol. Eds.) *Studies in Writing: Vol. 3. The Cognitive Demands of Writing*. (pp. 13-23). Amsterdam: Amsterdam University Press.
- Foulin, J.-N. (1998). To what extent does pause location predict pause duration in adults' and children's writing? *Cahiers de Psychologie Cognitive*, 17, 601-620.
- Gass, S. M. (2008). Second Language Acquisition: An Introductory Course. Routledge.
- Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 674–691.
- Gonzalez, C., Anderson, J., Culmer, P., Burke, M.R., & Mon-Williams, M. (2011). Is tracing or copying better when learning to reproduce a pattern? *Experimental Brain Research*, Vol. 208 (3) 459-465.
- Goodman, K.S. (1967). Reading: a psycholinguistic guessing game. *Journal of the Reading Specialist*, 4, 126-35.
- Gough, P. B. (1972). One second of reading. In J. F. Kavanagh & G. Mattingly (Eds.), *Language by ear and by eye* (pp. 331-358). Cambridge, MA: MIT Press.
- Grabe, W. & Stoller, F.L. (2002). *Teaching and researching reading* London: Longman.
- Grabowski, J. (2008). The internal structure of university students' keyboard skills. *Journal of Writing Research*, 1, 27–52.
- Grabowski, J. (2009). Speaking, writing, and memory span in children: Output modality affects cognitive performance. *International Journal of Psychology*, in press.
- Grabowski, J., Blabusch, C. & Lorenz, Th. (2007). Welche Schreibkompetenz? – Handschrift und Tastatur in der Hauptschule. [Which writing competence? – Handwriting and keyboard in German Hauptschule.]. In M.Becker-Mrotzek & K.Schindler (Eds.), *Texte schreiben*. (pp. 41–61). Kölner Beiträge zur Sprachdidaktik, 5. Köln: Gilles & Francke.
- Grabowski, J., Weinzierl, C., & Schmitt, M. (2010). Second and fourth graders' copying ability: From graphical to linguistic processing. *Journal of Research in Reading*, 33, 39-53.

- Graham, S., & Harris, K. R. (2000). The role of self-regulation and transcription skills in writing and writing development. *Educational Psychologist*, 35(1), 3–12.
- Graham, S., Berninger, V. W., Abbot, R. D., Abbot, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school children: A new methodological approach. *Journal of Educational Psychology*, 89, 170–182.
- Grainger, J. & Whitney, C. (2004) Does the huamn mnid raed wrods as a wlohe? Trends Cognition. Sci. 8, 58–59
- Hayes, J.R. & Flower, L. (1980). 'Identifying the organisation of the writing process'. In L. W. Gregg & E. R Steinberg (eds.). *Cognitive processes in writing*, 3-30. Hillsdale, NJ: Larence Erlbaum Associates.
- Haynes, M., & Carr, T. H. (1990). Writing system background and second language reading: A component skills analysis of English reading by native speaker-readers of Chinese. In T. H. Carr (Ed.), *Reading and its development: Component skills approaches* (pp. 375–421). San Diego, CA: Academic Press.
- Herrick, V.E. and Okada, N. (1963). The present scene: Practices in the teaching of handwriting in the United States. In V.E. Herrick (Ed.), *New horizons for research in handwriting*, University of Wisconsin Press, Madison, WI
- Hino, Y., & Lupker, S. J. (1998). The effects of word frequency for Japanese kana and kanji words in naming and lexical decisions: Can the dual-route model save the lexical selection account. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1431–1453.
- Hollerbach, J.M. (1981). An oscillation theory of handwriting. *Biological Cybernetics*, 39, pp. 139–156
- Holm, A., & Dodd, B. (1996). The effect of first written language on the acquisition of English literacy. *Cognition*, 59(2), 119–147.
- Hudson, R.F., Pullen, P.C., Lane, H.B., & Torgesen, J.K. (2009). The complex nature of reading fluency: A multidimensional view. *Reading & Writing Quarterly*, 25(1), 4-32
- Hue, C. W. (1992). Recognition processing in character naming. In H. C. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 93–107). Amsterdam: North-Holland.
- Jackson, N. E., Lu, W. H., & Ju, D. (1994). Reading Chinese and reading English: Similarities, differences, and second-language reading. In V. W. Berninger (Ed.), *The varieties of orthographic knowledge 1: Theoretical and developmental issues. Neuropsychology and cognition* (Vol. 8, pp. 73–109). Norwell, MA: Kluwer Academic.

- James, W. (1890) *The principles of psychology*. Henry Holt.
- Johnson, P. (1982). Effects on reading comprehension of building background knowledge. *TESOL Quarterly* 16(4):503-516.
- Just, M.A. & Carpenter, P.A. (1980). A theory of reading: from eye fixations to comprehension. *Psychology Review* 87:329–354.
- Kandel, S. & Valdois, S. (2006). Syllables as functional units in a copying task: A visuo-orthographic and graphomotor approach. *Language and Cognitive Processes*.
- Kandel, S. & Valdois, S. (2006a). Syllables as functional units in a copy task. *Language and Cognitive Processes*, 21, 432–452.
- Kandel, S. & Valdois, S. (2006b). French and Spanish-speaking children use different visual and motor units during spelling acquisition. *Language and Cognitive Processes*, 21, 531–561.
- Kandel, S., Herault, L., Grosjacques, G., Lambert, E., Fayol, M. (2009). Orthographic vs. phonologic syllables in handwriting production. *Cognition* 110. 3: 440-444.
- Kandel, S., Álvarez, C.J., Vallée, N. (2006). Syllables as processing units in handwriting production. *Journal of Experimental Psychology: Human Perception and Performance* 32. 1: 18-31.
- Kao, H. S. R., Shek, D. T. L., & Lee, E. S. P. (1983). Control modes and task complexity in tracing and handwriting performance. *Acta Psychologica*, 54, 69e77.
- Kao, H.S.R., Mak, P.H. & Lam P.W. (1986). Handwriting pressure: Effects of task complexity, control mode and orthographic difference. In H.S.R. Kao, G.P. van Galen, R. Hoosain (Eds.), *Graphonomics: Contemporary research in handwriting*, North-Holland, Amsterdam
- Kao, H.S.R., van Galen, G.P. & Hoosain, R. (1986). *Graphonomics. Contemporary research in handwriting*, North-Holland, Amsterdam.
- Kellogg, R.T. (1998). Components of Working Memory in Text Production. In G. Rijlaarsdam & E. Esperet (Series Ed.) & G. Jefferey (Vol. Eds.) *Studies in Writing: Vol. 3. The Cognitive Demands of Writing*. (pp. 43-62). Amsterdam: Amsterdam University Press.
- Khalidieh, S.A. (2001). The Relationship between Knowledge of Ibraab, Lexical Knowledge, and Reading Comprehension of Non-native Readers of Arabic. *The Modern Language Journal*, Volume 85 (3), 416-431.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production.

*Psychological Review*, 85, 363–394.

Kirk, U. (1980). Learning to copy letters: a cognitive rule-governed task. *Elementary School J* 81:28–33

Kluitmann, S. (2008). Testing English as a foreign language. (Thesis). Available from [http://www.sprachenmarkt.de/fileadmin/sprachenmarkt/wissen\\_images/Wissenschaftliche\\_Arbeit\\_Sebastian\\_Kluitmann\\_TOEIC-KMK.pdf](http://www.sprachenmarkt.de/fileadmin/sprachenmarkt/wissen_images/Wissenschaftliche_Arbeit_Sebastian_Kluitmann_TOEIC-KMK.pdf)

Koda, K. (1989). The effects of transferred vocabulary knowledge on the development of L2 reading proficiency. *Foreign Language Annals*, 22, 529-540.

Koda, K. (1999). Developing L2 intraword orthographic sensitivity and decoding skills. *Modern Language Journal*, 83(1), 51–64.

Koda, K. (2000). Cross-linguistic variations in L2 morphological awareness. *Applied Psycholinguistics*, 21, 297–320.

LaBerge, D., & Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293-323.

Lado, R. (1965). Memory span as a factor in second language learning. *IRAL*, 3, 2, pp. 123-29

Lambert, E., Kandel, S., Fayol, M., & Esperet, E. (2008). The effect of the number of syllables when writing poly-syllabic words. *Reading & Writing: An Interdisciplinary Journal*, 21, 859–883.

Laufer, B. & Goldstein, Z. (2004). Testing Vocabulary Knowledge: Size, Strength, and Computer Adaptiveness. *Language Learning* 54: 469-523

Laufer, B., & Nation, P. (1995). Vocabulary size & use: Lexical richness in L2 written productions. *Applied Linguistics* 16 (3), 307-322.

Laufer, B., Elder, C., Hill, K. & Congdon, P. (2004). Size and strength: Do we need both to measure vocabulary knowledge? *Language Testing* 21: 202-226

Leijten, M., & Van Waes, L. (2005). Writing with speech recognition: The adaptation process of professional writers with and without dictating experience. *Interacting With Computers*, 17, 736-772.

Leijten, M., & Van Waes, L. (2006). Inputlog: New perspectives on the logging of on-line writing processes in a Windows environment. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer key-stroke logging: Methods and applications* (pp. 73–93). Oxford: Elsevier.

Lewis, R. (1993). An Architecturally-based Theory of Human Sentence Processing. PhD Thesis, Carnegie

Mellon University, Pittsburgh, PA.

- Lindgren, E., & Sullivan, K. P. H. (2006). Analysing on-line revision. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer key-stroke logging: Methods and applications* (pp. 157–188). Oxford: Elsevier.
- Longcamp, M., Anton, J.L. & Velay, J.L. (2003). Visual presentation of single letters activates a pre-motor area involved in writing. *Neuroimage*, 19, 1492–1500.
- Longcamp, M., Boucard, C., Gilhodes, J.C. & Velay, J.C. (2006) Remembering the orientation of newly learned characters depends on the associated writing knowledge: a comparison between handwriting and typing. *Hum Mov Sci* 25:646–656
- Longcamp, M., Boucard, C., Gilhodes, J.C., Anton, J.L., Roth, M., Nazarian, B. & Velay JL (2008). Learning through hand- or typewriting influences visual recognition of new graphic shapes: behavioral and functional imaging evidence. *J CognNeurosci* 20:802–815
- Longcamp, M., Zerbato-Poudou, M.T. & Velay, J.L. (2005). The influence of writing practice on letter recognition in preschool children: a comparison between handwriting and typing. *ActaPsychol* 119:67–79
- Lorenz, T. & Grabowski, J. (2009). Handschrift oder Tastatur in der Hauptschule. Ökonomie und Mythen von Schriftmedien. [Handwriting or keyboard in German Hauptschule. Economy and myths of writing media.]. In K.Schneider, G.Schwab & M.Weingardt (Eds.), *Hauptschulforschung konkret*. (pp. 147–157). Baltmannsweiler: Schneider Verlag Hohengehren.
- M. Peck, E.N. Askov & S.H. Fairchild. (1980). Another decade of research in handwriting: Progress and Prospect in the 1970s. *Journal of Educational Research*, 73 (1980), pp. 283–298
- Marshall, J.C. & Newcombe, F. (1973). Patterns of paralexia: A psycholinguistic approach. *Journal of Psycholinguistic Research*, 2, 175-199.
- Matsuhashi, A. (1981). Pausing and planning: The tempo of written discourse production. *Research in the Teaching of English*, 15(2), 113-134.
- Matsuhashi, A. (1982). Explorations in the real-time production of written discourse. In: M. Nystrand (Ed.), *What writers know. The language, process, and structure of written discourse* (pp. 269-290). New York: Academic Press, Inc.
- Matsuhashi, A. (1987). Revising the plan and altering the text. In: A. Matsuhashi (Ed.), *Writing in real time – modelling production processes* (pp. 197-223). Norwood, NJ: Ablex Publishing company.
- McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational*

- Psychology Review*, 8, 299–325.
- McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist*, 35(1), 13–23.
- Meara, P. & Fitzpatrick, T. (2000). Lex30: An improved method of assessing productive vocabulary in an L2. *System* 28,1, 19-30.
- Meara, P. (1992). EFL vocabulary tests. Unpublished manuscript, Centre for Applied Language Studies, University College, Swansea, UK.
- Meara, P. (1993). Tintin and the world service: A look at lexical environments. *IATEFL: Annual Conference Report*, 32-37.
- Medwell, J. & Wray, D. (2007). Handwriting: What do know and what do we need to know? Literacy, Vol. 41, No. 1. Handwriting (pp. 10-15). Blackwell Publishing.
- Mehler, J. (1963). Some effects of grammatical transformations on the recall of English sentences. *Journal of Verbal Learning and Verbal Behavior*, 4, 748-762.
- Miller, G. A. (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63:81–97.
- Miller, G.A. (1962). Some psychological studies of grammar. *American Psychologist*, 11, 748-762.
- Miller, G.A., & Mc Kean, K. (1964). A chronometric study of some relations between sentences. *Quarterly Journal of Experimental Psychology*, 16, Part 4, 297-308.
- Moats, L. C. (2005). How spelling supports reading. *American Educator*, Winter 2005/06, 12-43.
- Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76(2), 165–178.
- Muljani, D., Koda, K., & Moates, D. R. (1998). The development of word recognition in a second language. *Applied Psycholinguistics*, 19(1), 99–113.
- Nation, I.S.P. & Beglar, D. (2007). A vocabulary size test. *The Language Teacher* 31, 7: 9-13.
- Nation, I.S.P. & Meara, P. (2002). Vocabulary. In N. Schmitt (ed.) *An Introduction to Applied Linguistics* Edward Arnold: 35-54.
- Nation, I.S.P. (2000). Vocabulary. In M. Byram (ed.) *Routledge Encyclopaedia of Language Teaching and Learning*. Routledge, London: 665-667.
- Nation, I.S.P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern*

*Language Review* 63, 1: 59-82.

Nation, I.S.P. (2007). The four strands. *Innovation in Language Learning and Teaching* 1, 1: 1-12.

Obaidellah, U.H. & Cheng, P.C.H. (2009). Graphical Production of Complex Abstract Diagrams: Drawing Out Chunks and Schemas. Proceedings of the Thirty First Annual Conference of the Cognitive Science Society. Amsterdam, Netherlands.: Cognitive Science Society.

Oxford University Press (2009). Oxford Practice Grammar, Diagnostic Test. Retrieved July 2009, from <http://www.oup.com/elt/global/products/practicegrammar/test/>.

Paris, S.G., Lawton, T.A., Turner, J.C., & Roth, J.L. (1991). A developmental perspective on standardised achievement testing. *Educational Researcher*, 20(5), 12-20.

Patterson, K. E. & Shewell, C. (1987). "Speak and spell: Dissociations and word class effects". In *The cognitive neuropsychology of language*, Edited by: Coltheart, M., Job, R. and Sartori, G. Hove, , UK: Lawrence Erlbaum Associates Ltd.

Pellicer-Sanchez, A. (2011). "Using eye-tracking methodology to measure L2 vocabulary acquisition from reading." Innovative approaches to language testing research [Conference]. BAAL TEASig 2011.

Pennington, M.C. (1999). 'Phonology in the context of communication and language learning', *Research Report*, Series 2. University of Luton.

Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. In C. Brown & P. Hagoot (Eds.), *The neurocognition of language* (pp. 167–208). New York: Oxford University Press.

Perfetti, C. A., & Zhang, S. (1995). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 24–33.

Perfetti, C.A. (1999). Comprehending written language: A blueprint of the reader. In C. Brown & P. Hagoot (Eds.), *The neurocognition of language* (pp. 167–208). New York: Oxford University Press.

Plaut, D.C., McClelland, J.L., Seidenberg, M.S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103, 56–115.

Power, M. J. (1986). A technique for measuring processing load during speech production. *Journal of Psycholinguistic Research*, 15, 371-382.

R. Plamondon, C.Y. Suen, M.L. Simner (Eds.). (1989). Computer recognition and human production of handwriting, World Scientific Publishing Co, Singapore



- R. Sassoon, I. Nimmo-Smith, A.M. Wing. (1986) An analysis of children's penholds. H.S.R. Kao, G.P. van Galen, R. Hoosain (Eds.), *Graphonomics: Contemporary research in handwriting*, North-Holland, Amsterdam (1986)
- Rawlinson, G. E. (1976) The significance of letter position in word recognition. Unpublished PhD Thesis, Psychology Department, University of Nottingham, Nottingham UK.
- Rawlinson, G. E. (1999). " Reibadaily", *New Scientist* (vol 162 issue 2188), May 1999, p.55.
- Rayner, K., & A. Pollatsek. 1989. *The psychology of reading*. New York: Prentice-Hall.
- Rayner, K., White, S. J., Johnson, R. L., & Liversedge, S. P. (2006). Raeding wrods with jubmled lettres: There is a cost. *Psychological Science*, 17, 192–193.
- Recht, D.R. & Leslie, L. (1988).Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*, 80:16-20.
- Rieben, L. & Saada-Robert, M. (1997). Relations between word-search strategies and word-copying strategies in children aged 5 to 6 years old. In C. A.Perfetti, L.Rieben & M.Fayol (Eds.), *Learning to spell: Research, theory, and practice across languages*. (pp. 295–318). Mahwah, NJ: Lawrence Erlbaum Associates.
- Rieben, L., Meyer, A. & Perregaux, C. (1991). Individual differences and lexical representations: How five 6-year-old children search for and copy word. L. Rieben, C. Perfetti, Editors , *Learning to read—Basic research and its implications*, Erlbaum, Hillsdale, NJ (1991), pp. 85–101.
- Rosenblum, S. (2005).Using the Alphabet Task to Differentiate between Proficient and Nonproficient Handwriters. In *Perceptual and Motor Skills* (pp. 629-639).
- Rubin, J. (1981). Study of Cognitive Processes in Second Language Learning1.*Applied linguistics*, 2(2), 117-131.
- Rubin, J. (1987). Learner strategies: Theoretical assumptions, research history and typology. *Learner strategies in language learning*, 15-30.
- Saada-Robert, M. & Rieben, L. (1993). Evolutions des stratégies d'écriture-copie et unités graphiques du français. *Etudes de Linguistique Appliquée*, 90, 84–96.
- Salah, S.M. (2008) The relationship between vocabulary knowledge and reading comprehension of authentic arabic texts (Unpublished master's thesis). Brigham Young University, Provo, USA.
- Saville-Troike, M. (2006). *Introducing second language acquisition*. Cambridge: Cambridge University

Press.

- Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. Neuman & D. Dickinson (Eds.), *Handbook for research in early literacy* (pp. 97–110). New York: Guilford Press.
- Schilperoord, J. (1996) It's about time. Temporal aspects of cognitive processes in text production. Amsterdam: Rodopi.
- Schilperoord, J. (2001). On the cognitive status of pauses in discourse production. In: G. Rijlaarsdam (Series ed.) & T. Olive & C.M. Levy (Vol. eds.), *Studies in Writing: Volume 10: Contemporary Tools and Techniques for Studying Writing*. Kluwer Academic Publishers: Netherlands.
- Schilperoord, J. (2003). On the cognitive status of pauses in discourse production. In T. Olive & C. M. Levy (Eds.), *Contemporary tools and techniques for studying writing* (pp. 61-88). Dordrecht: Kluwer.
- Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language*, 37, 118–139.
- Sears, C. R., Sharp, C. R., & Lupker, S. J. (2006). Is there a neighbourhood frequency effect in English? Evidence from reading and lexical decision. *Journal of Experimental Psychology: Human Perception & Performance*, 32, 1040–1062.
- Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition*, 19, 1–30.
- Seidenberg, M. S., Waters, G. S., & Barnes, M. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior*, 23, 383–404.
- Severinson Eklundh, K., & Kollberg, P. (1996a). A computer tool and framework for analysing on-line revisions. In C.M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 163–188). Mahwah, NJ: Lawrence Erlbaum Associates.
- Severinson Eklundh, K., & Kollberg, P. (1996b). Computer tools for tracing the writing process: From keystroke records to S-notation. In G. Rijlaarsdam, H. van den Bergh, & M. Couzijn, M. (Eds.), *Current Research in writing: Theories, models and methodology* (pp. 526–541). Amsterdam: Amsterdam University Press.
- Shu, H., & Anderson, R. C. (1997). Role of radical awareness in the character and word acquisition of Chinese children. *Reading Research Quarterly*, 32(1), 78–89.
- Shu, H., Anderson, R. C., & Wu, N. N. (2000). Phonetic awareness: Knowledge of orthography-phonology

- relationships in the character acquisition of Chinese children. *Journal of Educational Psychology*, 92(1), 56–62.
- Simpson, S., & Torrance, M. (2007). EyeWrite (Version 5.1). Osgoode, ON: SR Research, and Nottingham, U.K.: Nottingham Trent University.
- Smith, F. (1971). Understanding reading. New York: Holt, Rinehart & Winston. 77, 4, 257-273.
- Søvik, N. (1975). Developmental cybernetics of handwriting and graphic behavior. Universitetsforlaget, Oslo
- Spelman Miller, K. & Sullivan, K. P. H. (2006). Keystroke logging: an Introduction. In G. Rijlaarsdam (Series Ed.) and K.P.H. Sullivan, & E. Lindgren. (Vol. Eds.), *Studies in Writing*, Vol.18, Computer Keystroke Logging: Methods and Applications. Oxford; Elsevier.
- Spelman Miller, K. (2000a). Academic writers on-line: Investigating pausing in the production of text. *Language Teaching Research*, 4(2), 123–148.
- Spelman Miller, K. (2000b). *Writing online: Temporal features of first and second language written text production*. Unpublished doctoral thesis. University of Reading, UK.
- Spelman Miller, K. (2006). The pausological study of written language production. In G. Rijlaarsdam (Series Ed.) and K.P.H. Sullivan, & E. Lindgren. (Vol. Eds.), *Studies in Writing*, Vol.18, Computer Keystroke Logging: Methods and Applications. Oxford; Elsevier.
- Stelmach, G.E.; Diggles, V.A.(1982) Control Theories in Motor Behavior. *Acta Psychologica*, Vol.50, No.1, pp.83-105(1)
- Stephanie, Y. H. M. (2009). Predictors of Spelling Development in Bilinguals: A Comparison of English L1 and Mandarin L1 Children.
- Stromqvist, S. & L. Malmsten. 1998. *Scriptlog pro 1.04: user's manual*. Department of Linguistics, University of Goteborg.
- Strömquist, S., & Ahlsén, E. (Eds.). (1999). *The process of writing: A progress report*. Göteborg, Sweden: Göteborg University, Department of Linguistics.
- Strömquist, S., & Karlsson, H. (2002). *ScriptLog for Windows: User's manual* (Tech. rep.). Lund: Lund University, and University College of Stavanger, Centre for Reading Research, Department of Linguistics.
- Strömquist, S., Holmqvist, K., Johansson, V., Karlsson, H., & Wengelin, Å. (2006). What keystroke logging

- can reveal about writing. In G. Rijlaarsdam (Series Ed.) and K. P. H. Sullivan & E. Lindgren (Eds.), *Computer key-stroke logging: Methods and applications* (pp. 45–71). Oxford: Elsevier.
- Süselbeck, G. (2003). Abschreiben: eine wichtige Arbeitstechnik. [Copying: an important working technique.]. *Grundschule*, 35 (2), 16–20.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development in Gass, S. and Madden, C. (eds.), *Input in second language acquisition* (pp. 235-256). Rowley, MA: Newbury House.
- Taft, M. (1991). *Reading and the mental lexicon*. Hove, England: Erlbaum.
- Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language & Cognitive processes*, 9, 271-294.
- Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading*, 10, 381–398.
- Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word recognition. *Journal of Memory and Language*, 26, 608–631.
- Taylor, B. M. (1979), Good and poor readers' recall of familiar and unfamiliar text. *Journal of Reading Behavior*, 11, 375-380.
- Taylor, L. "Using eye-tracking technology to research L2 reading." Innovative approaches to language testing research [Conference]. BAAL TEASig 2011. 18 November 2011.
- Teulings, H.L. & Maarse F.J. (1984). Digital recording and processing of handwriting movements. *Human Movement Science*, 3, pp. 193–217
- Teulings, H.L. & Thomassen, A.J.W.M. (1979). Computer-aided analysis of handwriting movement. *Visible Language*, 13, pp. 219–231
- Thomassen, A.J.W.M., Keuss, P.J.G., van Galen, G.P. (1984) *Motor aspects of handwriting: Approaches to movement in graphic behavior*. North-Holland, Amsterdam
- Torrance, M. & Jeffery, G. (1998). *Writing Processes and Cognitive Demands*. In G. Rijlaarsdam & E. Esperet (Series Ed.) & G. Jefferey (Vol. Eds.) *Studies in Writing: Vol. 3. The Cognitive Demands of Writing*. (pp. 1-12). Amsterdam: Amsterdam University Press.
- Torrance, Mark; Galbraith, David (2006) *The processing demands of writing*. In. MacArthur, Charles; Graham, Steve; Fitzgerald, Jill (eds.) *Handbook of writing research*. New York: The Guilford Press, 67-

80.

Quinlan, T. (2010) ETS – SigWriting Conference Paper Presentation

Van der Gon Denier, J.J. Thuring, J.Ph. Thuring. (1985). The guiding of human writing movements. *Kybernetik*, 2, pp. 145–148

Van Galen, G. P., (1991). Handwriting: issues for a psychomotor theory. *Hum. Mov. Sci.* 10, 165–191.

Van Galen, G.P. & Teulings, H.L. (1983). The independent monitoring of form and scale factors in handwriting. *Acta Psychologica*, 54, pp. 9–22

Van Galen, G.P. (1990). Phonological and motoric demands in handwriting: Evidence for discrete transmission of information. *Acta Psychologica*, 74 , pp. 259–275

Van Galen, G.P., Meulenbroek, R.G.J. & Hylkema, H. (1986). On the simultaneous processing of words, letters and strokes in handwriting: Evidence for a mixed linear and parallel model. In H.S.R. Kao, G.P. van Galen, R. Hoosain (Eds.), *Graphonomics: Contemporary research in handwriting*, North-Holland, Amsterdam

Van Galen, G.P., Smyth, M.M., Meulenbroek, R.G.J. & Hylkema, H. (1989). The role of short-term memory and the motor buffer in handwriting under visual and non-visual guidance. In R. Plamondon, C.Y. Suen, M.L. Simner (Eds.), *Computer recognition and human production of handwriting*, World Scientific Publishing Co, Singapore

Van Genuchten, E. (2009). Pause Length Extraction Tool (PLET), Computer software. University of Sussex. Brighton, UK.

Van Genuchten, E., & Cheng, P. C.-H. (2010). Temporal chunk signal reflecting five hierarchical levels in writing sentences. In S. Ohlsson & R. Catrambone(Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. (pp. 1922-1927). Austin, TX: Cognitive Science Society.

Van Genuchten, E., Cheng, P. C.-H., Leseman, P. P. M., & Messer, M. H. (2009). Missing working memory deficit in dyslexia: Children writing from memory. In N. A. Taatgen, & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 1674-1679). Austin, TX: Cognitive Science Society.

Van Orden, G. (1987). A ROWS is a ROSE: Spelling, sound and reading. *Memory and Cognition*, 15: 181–198.

Van Orden, G. C. ,& Goldinger, S. D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and*

*Performance*, 20, 1269-1291.

- VanOrden, G.C. & Goldinger, S.D. (1996). Phonologic mediation in skilled and dyslexic reading. In C. Chase, G. Rosen, & G. Sherman (Eds.), *Developmental dyslexia: Neuro, cognitive, and genetic mechanisms* (pp. 185–223). Timonium, MA: York Press.
- Velan, H., & Frost, R. (2007). Cambridge University vs. Hebrew University: The impact of letter transposition on reading English and Hebrew. *Psychonomic Bulletin & Review*, 14, 913–918.
- Verhoeven, L., Schreuder, R. & Haarman, V. (2006). Prefix identification in the reading of Dutch bisyllabic words. *Reading and Writing*, 19 (2006), pp. 651–668
- Verhoeven, L., Schreuder, R., Baayen, R.H. (2006) Learnability of graphotactic rules in visual word identification. *Learning and Instruction*, 16 (6), pp. 538-548.
- VocabProfile (2006) VP English Version 3 <http://www.lexutor.ca/vp/>
- Vredenburg, J. & Koster, W.G. (1971). Analysis and synthesis of handwriting. *Philips Technical Review*, 32, pp. 73–78
- Wang, M., & Geva, E. (2003). Spelling performance of Chinese ESL children: Lexical and visual-orthographic processes. *Applied Psycholinguistics*, 24, 1–25.
- Wang, M., & Koda, K. (2005). Commonalities and differences in word identification skills among learners of English as a second language. *Language Learning*, 57(1), 201-222.
- Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and non-alphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition*, 87, 129–149.
- Weir, C.J. and H. Khalifa. (2008). 'A cognitive processing approach towards defining reading comprehension'. *Cambridge ESOL: Research Notes* 31: 2-10
- Weinzierl, C., Grabowski, J., & Schmitt, M. (2012). Copying ability in primary schools: Working memory when copying texts. *Learning to Write Effectively: Current Trends in European Research*, Prepublication Draft (26-28).
- Wengelin, Å. (2006). Examining pauses in writing: theory, methods and empirical data. In G. Rijlaarsdam (Series Ed.) and K.P.H. Sullivan, & E. Lindgren. (Vol. Eds.), *Studies in Writing*, Vol.18, *Computer Keystroke Logging: Methods and Applications* (pp. 107-130). Oxford; Elsevier.
- Wengelin, Å. (2007). The word-level focus in text production by adults with reading and writing difficulties. In Rijlaarsdam, G. (Series Ed.) and M. Torrance, L. van Waes & D. Galbraith (Volume Eds.), *Writing and Cognition: Research and Applications* (Studies in Writing, Vol.20, pp. 67-82).

Amsterdam: Elsevier.

- Wengelin, A., Torrance, M., Holmqvist, K., Simpson, S., Galbraith, D., Johansson, V., Johansson, R. (2009) Combined eyetracking and keystroke-logging methods for studying cognitive processes in text production. *Behavior Research Methods Journal*, 41 (2), 337-351.
- Wengelin, A°. (2002). Text writing in adults with reading and writing difficulties. Gothenburg Monographs in Linguistics 20.
- Wharton, G. (2000). Language learning strategy use of bilingual foreign language learners in Singapore. *Language learning*, 50(2), 203-243.
- White, S. J. (2008). Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance*, 34: 205–223.
- Wing, A.M. (1978). Response timing in handwriting. In G.E. Stelmach (Ed.), *Information processing in motor control and learning*, Academic Press, New York
- Wittrock, M.C. (1974). Learning as a generative process. *Educational Psychologist*, 11, 87-95.
- Wittrock, M.C., Marks, C.B., & Doctorow, M.J. (1975). Reading as a generative process. *Journal of Educational Psychology*, 67, 484-489.
- Zhang, F. & Yin, P. (2009). A study of pronunciation problems of English learners in China. *Asian Social Science* 5, 141-146